5-2017

Proactive and reactive strategies to handle surges in urban crowds

Jiali DU
Singapore Management University, jiali.du.2012@phdis.smu.edu.sg

PROACTIVE AND REACTIVE STRATEGIES TO HANDLE SURGES IN URBAN CROWDS

JIALI DU

SINGAPORE MANAGEMENT UNIVERSITY

2017

Proactive and Reactive Strategies to Handle Surges in Urban Crowds

by

Jiali Du

Submitted to School of Information Systems in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Information Systems

Dissertation Committee:

Shih-Fen Cheng (Supervisor/Chair)
Associate Professor
Singapore Management University

Hoong Chuin Lau (Co-supervisor)
Professor
Singapore Management University

Hai Wang
Assistant Professor
Singapore Management University

Erika Fille Tupas Legara
Scientist
Institute of High Performance Computing, A Star

Singapore Management University
2017

Proactive and Reactive Strategies to Handle Surges in Urban Crowds

Jiali Du

Abstract

Most urban infrastructures are built to cater a planned capacity, yet surges in usage do happen in times (can be either expected or unexpected), and this has long been a major challenge for urban planner. In this thesis, I propose to study approaches handle surges in urban crowd movement. In particular, the surges in demand studied are limited to situations where a large crowd of commuters/visitors gather in a small vicinity, and I am concerned with their movements both within the vicinity and out of the vicinity (the egress from the vicinity). Significant crowd build-ups and congestions can be observed in a number of cases I studied, and when capacity expansion is not a viable strategy (either because of budget or physical constraints), smoothing these demand surges would be the only practical solution. To handle such demand surges in urban crowds, we can either:

1. Distribute demands temporally: by slowing down the flow rate of incoming demands to the congested region through providing incentives or distractions.
2. Distribute demands spatially: by redirecting overflowing demands to other parts of network where spare capacities are still available. This might require additional investment in establishing complementary connection service.

My thesis aims at proposing computationally efficient strategies to tackle these issues. The first strategy targets on distributing demands temporally in a proactive way. In other words, this strategy is designed to prevent demand peaks from forming by slowing down crowd from congregating to areas of concern. As an example, I propose to study the strategy of crowd management in the context of theme park; in particular, the delay of flow rate towards congested areas is achieved by providing distractions (or incentives). It is demonstrated that crowd build-ups can be mitigated

by utilizing this strategy. However, it might not always be possible to delay the crowd movement. For example, after major sports events that end late, most of crowd would just want to leave the stadium and reach home as soon as possible, and they will not slow down their egress pace, regardless of distractions/incentives. In these cases, I propose to study the use of the second strategy, which distributes crowds spatially to other parts of network so as to avoid clogging the vicinity that is closest to the demand node. More specifically, I propose to provide parallel services complementing existing ones so that more commuters can leave overcrowded areas and start their trips from other less crowded nodes. Consequently, there should be much fewer people queuing for services at the origin node.

Contents

List of Figures

List of Tables

List of Notations

Variable	**Definition** (Chapter 3)
$p_{u,v}^t$	Transition probability between nodes u and v at time step t
$n_u^{t,d}$	Number of agents in node u at time t on day d
s_u	Service rate (number of agents serviced in one time step) of node u
D	Set of cascades or days on which observations about **n** are made
$x_{u,v}^{t,d}$	Number of agents moving from node u to node v at time t on day d
$l_{u,v}^{t,k}$	Binary variable that is set to 1 if side show of type k is placed on edge (u, v) at time t
β^k	Percentage of diffusion attracted by side show of type k

Variable	**Definition** (Chapter 4 & Chapter 5)
N	Node set with subset N_k representing nodes serve passengers in group k and N_r denote node set included in route r
R	Route set with subset R_k representing available routes for group k
K	Passenger groups, with demand d_k and destination π_k for each group k
Q	Capacity of transportation services
F	Maximum round of services deployed for a route
B	Budget given to deploy the service
δ	Travel time, including δ^r for route r and $\delta_{i,j}$ for arc (i, j)
$\boldsymbol{c}$	Vector denoting the cost of deployment
$\alpha_{i,j}^r$	Set to 1 if route r covers arc (i, j)
w_k^r	Set to 1 if demand k is served by route r
g_i^r	Set to 1 if station i is included in route r
$\beta_{i,j}$	The probability of passengers going from location i to j
$\underline{q}, \overline{q}$	Lower and upper bound of bike station capacity

Acknowledgement

First and foremost, I would like to express my appreciation to my supervisors Prof. Cheng Shih-Fen and Prof. Lau Hoong Chuin for their continuous support of my PhD study. I would like to thank you for supervising my research, making me grow to be a data scientist from a junior one. I appreciate all your contributions of time, ideas, and funding to support my study. Over the years, you have provided valuable and kind suggestions and helped on both research and my life, thank you for all of these.

Besides my supervisors, I would like to thank the rest of my thesis committee: Prof. Wang Hai and Dr. Erika Fille Tupas Legara, thank you for devoting your time and efforts and providing insightful comments on my thesis.

I would like express my thanks to Prof. Pradeep Varakantham, who spent lots of time on guiding me how to do research when I was a junior researcher. Thank you for your patience, hard work and kindness along the way. Thanks to Prof. Akshat Kumar, who spent holidays helping me improve my work and addressed very smart suggestions.

Dozens of people have helped me at SMU during the 5 years time. Thanks to my friend Deng Xiaoxia and her families for taking care of me and offering lots of help. Thanks to my office mates and friends: Du Juan and Gong Wei, for growing up with me together and making me a better person. Thanks to Nikita Jaiman for being a warm-hearted friend and help me overcome difficulties along the way. Thanks to Fu Na for giving advises for my life as a senior. Thanks to Prof. Lim Ee Peng, who makes LARC a warm family where I stayed for my PhD time. Thanks to Prof.

Robert Kauffman, who is an amiable elderly and I wish you healthy and happy. Thanks to Koh Yong Fong, for helping me and encouraging me in many ways when I was at a bad time. Thanks to Ong Chew Hong and Seow Pei Huan who assist my PhD study along the way. Thank you to the admin team at LARC, for continueously solve lots of trivial problems for me. Thanks to all the staffs, colleagues and friends at SMU, without all of you, I cannot make my 5 years happy and enriching.

Lastly, I would like to thank my families. My great husband Tang Jing, thank you for accompanying and helping me grow up from depression. It is not possible for me to insist for this long without you. Thanks to my parents for understanding and supporting their only child to be far away from home for nearly 10 years since 2008. I love you all and miss you all.

Publications

Publications based on the dissertation.

1. Jiali Du, Pradeep Varakantham, Akshat Kumar and Shih-Fen Cheng, *Learning and Controlling Network Diffusion in Dependent Cascade Models*, In Web Intelligence and Intelligent Agent Technology (WI-IAT-15). (Chapter 3).

2. Jiali Du, Akshat Kumar and Pradeep Varakantham, *On understanding diffusion dynamics of patrons at a theme park*, *Extended Abstract*, In International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-14). (Chapter 3).

3. Jiali Du, Shih-Fen Cheng and Hoong Chuin Lau, *Designing Bus Transit Services for Routine Crowd Situations at Large Event Venues*, In International Conference on Computational Logistics (ICCL-15). (Chapter 4).

Working papers based on the dissertation.

1. Jiali Du, Shih-Fen Cheng and Hoong Chuin Lau, *Adding Capacity without Building Capacity: A Spatial Redistribution Approach*, (*to be submitted to* EURO Journal on Transportation and Logistics). (Chapter 4).

2. Jiali Du, Shih-Fen Cheng and Hoong Chuin Lau, *Optimizing the Portfolio of Transport Services for Last-mile Commute*, (*to be submitted*). (Chapter 5).

Other publications during Ph.D. study.

1. Shih-Fen Cheng, Larry Lin, Jiali Du, Hoong Chuin Lau, Pradeep Varakantham, *An agent-based simulation approach to experience management in theme parks*, In Winter Simulations Conference (WSC-13).

2. Shih-Fen Cheng, Larry Lin, Jiali Du, Hoong Chuin Lau, and Pradeep Varakantham, *An agent-based simulation approach to experience management in theme parks*, In International Workshop on Multi-Agent-Based Simulation (MABS-13).

3. Na Fu, Jonathan Chase, Jiali Du, Truc Viet Le and Hoong Chuin Lau, *Data-driven Law Enforcement Staffing and Allocation with Quality of Service Guarantees*, *Accepted by* Machine Learning, Optimization, and Big Data (MOD-17).

Chapter 1

Introduction

1.1 Motivation

For the past few decades, we have witnessed unprecedented growth in the degree of urbanization. The percentage of global population living in cities has already reached 54% in 2014, and is expected to reach two-third by 2050[1]. The increasing population leads to higher complexity in urban city management. In particular, most existing infrastructures are built to cater a planned capacity, yet the surges in usage do happen in times (can be either expected or unexpected), and this has been a major challenge for urban planner.

Challenges include but not limited to: (1). Reducing the congestions in popular theme park to prevent passengers from waiting for too long; (2). Managing the traffic support for large crowds after major events to reduce the congestion; (3). Balancing the supply and demand for taxis service during peak hours so that to reduce the inconvenience for passengers; (4). Adjusting the toll price to control the traffic flow on express road and so on.

To tackle the demand surge issues for urban crowds, I propose to study approaches which can handle surges in urban crowd movement. Surge of demand should be quantified by analysis on the data set and the proposed strategy is ex-

[1] http://www.un.org/en/development/desa/news/population/world-urbanization-prospects-2014.html

pected to be practical and not too costly.

Demand surge generally describes the phenomenon where excessive demand occurs suddenly but the supply is not prepared to respond quickly. Specifically, the surges in demand studied in this thesis are limited to situations where a large crowd of commuters/visitors gathering in a small vicinity, and I am concerned with their movements both within the vicinity and out of the vicinity (the egress from the vicinity). Vicinity studied in this thesis include national stadium, major MRT stations, popular theme park. Surges studied are triggered by the large event, morning evening peak hours and popular attractions.

Demand surges cause inconvenience for people and leads to various problems. For instance, in a large theme park of Singapore, historical data shows an average waiting time of over 80 minutes for the popular attractions during peak hours[7], which heavily impacts visitors' experience. Figure 1.1 provides details of waiting time for different attractions at different time of a day, from which we observe that attraction T has a long queue (over 80 mins) in early morning. Another example

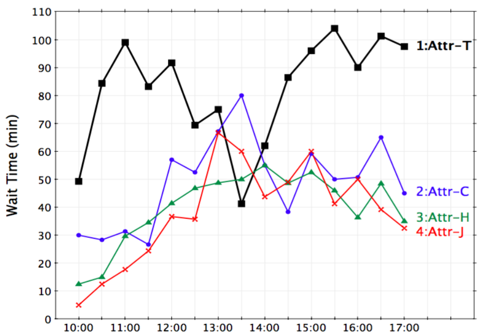


Figure 1.1: Historical waiting time in theme park [7]

addresses the problem of egress when huge number of people gather at the facilities such as stadium or conventional centers during major events. Such demands are going to overwhelm the existing infrastructures significantly. Peak-hour traffic congestion in growing metropolitan areas is another concern for urban planners. In a business park, the majority of people seeking to move in during morning rush hours, which causes heavy congestions in MRT stations nearby.

Having the necessary resources to respond to surging demand in a timely manner is therefore, very important and mechanisms must be proposed to handle this efficiently in order to enhance the experience. To handle such demand surges in urban crowds, we can either:

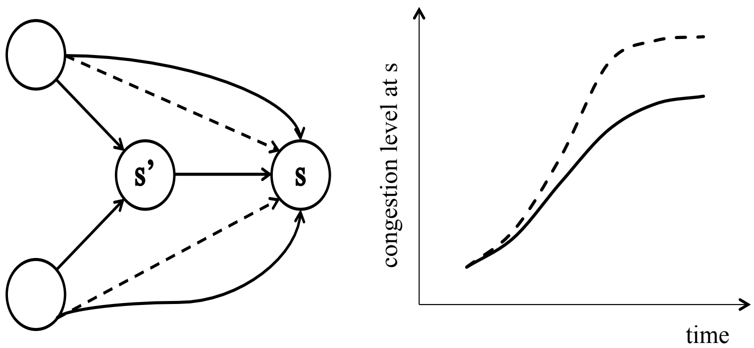
(1). **Distribute demands temporally** by slowing down the flow rate of incoming demands to the congested region through providing incentives or distractions.

(2). **Distribute demands spatially** by redirecting overflowing demands to other parts of network where spare capacities are still available. This might require additional investment in establishing complementary connection service.

My thesis aims at proposing computationally efficient strategies to tackle these issues.

1.2 Research Objectives

We discuss two strategies in this thesis. The first strategy targets on distributing demands temporally in a proactive way. In other words, this strategy is designed to prevent demand peaks from forming by slowing down crowd from congregating to areas of concern.

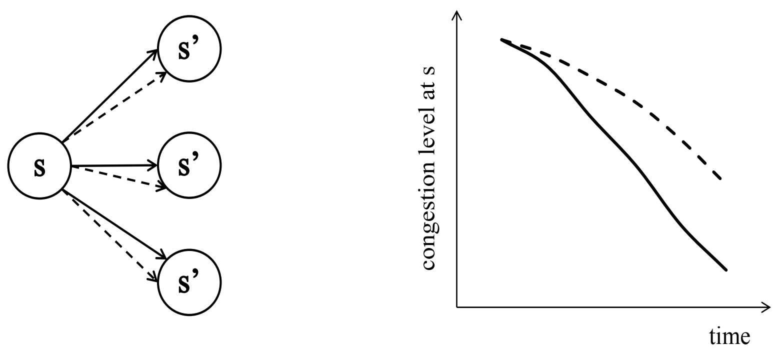


(a) Movement of flow (b) Congestion level at node s

Figure 1.2: Illustration of the first strategy – using s' to slow down the crowds and hence reduce the congestion level

Figure 1.2 helps explain the idea. Formally, recognizing the high likelihood of flow moving towards a busy node, s, in a short time, it is possible to make people take detour to a leisure node, s', before they reach s. The dotted line with arrows in Figure 1.2(a) represents the original direction of flow and the solid line denotes the actual flow after diverting. Remaining at s' takes some time and helps slow down the increasing speed of the flow to s. As can be seen in Figure 1.2(b), without any intervention, the congestion level at node s quickly increases over time shown with the dotted line. After deploying the strategy, the increasing speed on the congestion level at node s is mitigated described by the solid line.

However, it might not always be possible to delay crowd movement. For example, after major sports events that end late, most would just want to leave the stadium and reach home as soon as possible, and they will not slow down their egress pace, regardless of distractions/incentives. In these cases, I propose the second strategy, which distributes crowds spatially to other part of network quickly so as to avoid clogging the node that is closest to the demand node.



(a) Movement of flow (b) Congestion level at node s

Figure 1.3: Illustration of the second strategy – adding capacity between s and s' to increase the dispersion rate of crowds hence reduce the congestion level

Formally, when observing large quantities of flow accumulated at node s, we target on accelerating the dispersal process of flow going from node s to its neighbors, s'. Refer to Figure 1.3(a), we use dotted lines to represent the original capacity between s and s' and the solid lines denote the additional capacity provided to increase the throughput between nodes s and s'. After putting in more spare capacities, the quantity of flow clogged at node s represented by solid line is reduced compared to the original one denoted by dotted line in Figure 1.3(b).

In the following sections, we use three real-world applications to explain how we apply the two strategies to handle the demand surges.

1.2.1 Learning and Controlling Network Diusion: A Temporal Redistribution Approach

In this section, I propose to study the first strategy of crowd management in the context of theme park. In particular, the delay of flow rate towards congested areas is achieved by providing distractions (or incentives). We model the visitors' mobility behaviors in the theme park as diffusion dynamics.

In general, diffusion processes describe how ideas, influence, and people spread over an underlying network, which for example may be a social network [27] or a transportation network [29]. Understanding diffusion dynamics is important as it helps predict and control the contagion spread in a network. The independent cascade (IC) model of [27] and its variants have been quite successful in modeling such a diffusion process over a network in a variety of domains [34],[28],[18].

We address several features of cascades in real-world that are not modeled in the existing IC model. For example, in many networks, it is not feasible to track the status of each entity in the cascade (e.g., tracking visitors moving in a theme park). Therefore, diffusion dynamics must be learned from the *aggregate* observed data for the underlying contagion. Some examples include finding the most probable path of migratory birds [13], understanding the evolution of traffic in a transportation network [8],[29], emotional contagion in a crowd evacuation scenario [54] and mobility pattern of visitors in a theme park. Therefore, we develop techniques based on mathematical optimization that can learn the underlying dynamics of the diffusion process using only aggregate data.

We incorporate realistic features, such as modeling queues at nodes in the underlying network (e.g., attractions in a theme park) while learning the diffusion dynamics based on aggregate data. We further augment the IC model with flow conservation that is required when modeling the diffusion process in problems such as traffic flow diffusion or visitor flow across theme park attractions. Incorporating such features results in a *dependent cascade* model, where the diffusion over the outgoing edges must satisfy flow conservation and is no longer independent for each edge.

Given the learned model of the diffusion dynamics, the next key problem we address is how to take decisions within a given budget to alter the dynamics of the underlying diffusion process to optimize some performance criterion. We are motivated by the problem of placing sideshows in an overcrowded theme park to reduce congestion at different attractions. Sideshows alter the underlying flow of visitors by attracting some fraction of visitors to themselves, thereby reducing congestion at main attractions in the theme park. When and where to place sideshows using limited resources is the key decision making problem we address. We show the research of this part in Chapter 3.

1.2.2 Bus Bridging in Post-event Crowd Diffusion: A Spatial Redistribution Approach

In this section, I propose to study the use of the second strategy, which distributes crowds spatially to other part of network quickly so as to avoid clogging the node that is closest to the demand node. More specifically, I propose to provide parallel services complementing existing ones so that more commuters can leave overcrowded areas and start their trips from other less crowded nodes.

In urban city, there is a growing number of large events such as concert, sports game and festivals. Operating the events is challenging as the surge of human traffic demand caused by the events impose a ultra-high traffic stress in transportation. Impacts include but not limited to slow road traffic, delay on public transportation, extremely long wait time for passengers and so on.

Our aims in this work are to provide a formal model and a computationally scalable approach for handling demand surges concentrating at a few highly congested origins. In particular, we propose a "spatial redistribution" approach, which optimally reroutes public transport passengers experiencing excessive wait times at known origins to other less congested parts of the public transport network. As these rerouting efforts are achieved mostly by providing complementary shuttle bus services linking pairs of existing public transport stations, the capacity of these connection services will be limited by the available budget. Due to this budget limit, and the fact that the number of service lines is exponential in the number of stops, a computationally efficient optimization approach will be necessary to overcome the combinatorial nature of the optimization problem.

To make our approach computationally feasible, we propose a two-phase framework which first decides which shuttle bus links should be established, followed by deciding the capacity for each established link. During the first phase, we implement a column generation procedure, which iteratively identifies new route that could satisfy more demands. The key idea of this phase is to identify a portfolio of additional connection services that could serve as many passengers as possible. Given the set of service lines to establish, the second phase will then assign available buses to individual service lines, minimizing the total travel time experienced by commuters while observing the budget constraint.

1.2.3 Integrated Bus and Bike Sharing Services for Last-mile Commute: A Spatial Redistribution Approach

In this section, I propose to study the use of the second strategy to distribute crowds spatially in the context of business district. In the main station of the business park, we observe a clear demand surge pattern during morning and evening peak hours on working days. Typically, I propose to provide bimodal transportation services between the demand node and passengers' destination nodes, which expands the transportation capacities in two dimensions.

Many foreign companies prefer to establish their business hub in world-class cities such as Singapore, which, encourages the local government to develop the areas for building business districts. As more and more business districts are planned and constructed, their locations are not limited to downtown but cover suburban areas as well. Transportation network for business district in downtown center is vibrant. However, regarding suburban areas, there is only limited number of MRT stations in the business park, which is not sufficient to serve all passengers heading towards the business park. During morning and evening peak hours, demand surge occurs at those major MRT stations.

Our goal for this work is to provide a formal model that computes the optimal portfolio over two different transportation services for passengers to travel from major MRT stations to their working offices. Specifically, we present a computationally efficient approach which comes out with the best portfolio assignment plan over a mixture of transportation modes: bus and bike sharing services. Adopting both bus and bike services provide benefits to operators: bus allows flexible transportation such that the schedule can be altered according to real-time demand patterns; bike sharing service is cheap to deploy. Integrating both services provide a way to increase transportation capacities in parallel between the major MRT station and destinations.

As bus and bike planning problems are two independent problems, we propose each model in the first phase. The fundamental idea behind bus model is to determine good bus routes which can serve a group of passengers along its way and deploy buses over each candidate routes so that to achieve minimal travel time. The main issues solved by bike sharing model is to determine where to construct bike storing stations and its relative capacity. In the meanwhile, the model also produces the bike-rebalancing plan. With above model proposed, in the second phase we connect them using systematic methods – Lagrangian Relaxation and Sub-gradient approach. In this way, we iteratively obtain good plans which save passengers' travel time.

1.3 Contributions

To summarize, we review the contributions in this dissertation mainly in two aspects.

Methodologically speaking, the major contributions are:

1. For distributing demands along the temporal dimension, I apply the Dependent Cascade (DC) model to model visitors transition behaviors, which augments existing Independent Cascade (IC) models by incorporating a real-world feature of (1). adding flow conservation constraints at each node; and (2). learning the diffusion dynamics with only aggregate observations.

2. For distributing demands along the spatial dimension, I provide the optimization-based approaches to identify bus routes and their operational policies for connection service. To scale up computational, a particular variant of column generation technique is applied: I identify bus routes that minimize total travel cost in the master problem; I identify the beneficial route in the pricing sub-problem.

3. For distributing demands along the spatial dimension during morning and evening peak hours, I provide bimodal transportation options including both bus and bike sharing services to expand the transportation capacity along the way. To connect those two independent models systematically, I apply Lagrangian Relaxation and Sub-gradient method so that can iteratively obtain the solutions.

Practically speaking, the major contributions are:

1. I use the theme park context to illustrate the first strategy. I apply the real-world historical wait time data to learn visitors transition probabilities and the accuracy is close to 80%. With the learned parameters, links that should be slowed down will be identified, and distractions such as side shows or parades will be introduced to avoid crowd build-ups for busy attractions. Close to 20%

average wait time reduction and up to 25% peak wait time reduction can be achieved by my approach.

2. I use the public transportation context to illustrate the second idea. I validate the model using a concession card dataset in Singapore (which contains all tap-in and tap-off records for all commuters). Numerical experiments are conducted under two conditions: In the normal situation where the existing infrastructures are working well, a reduction of 18.8% in the total travel time is achieved; under the disrupted case where certain links are shut down, my approach is comparable to one of the state-of-art methods found in the literature (despite the fact that the method is not designed for the disrupted case).

3. I further use the bus services in the context of a business park to illustrate the second idea. I validate the model by comparing our bus service and mixed transportation options to the existing deployment addressed by the operators. Results indicate we generate better bus routes than the current ones and we save passengers' average travel time with bike services incorporated.

Generalization: the extension of research findings and strategies proposed in my thesis is generalizable due to following reasons:

1. The data set we collected are either from urban planners or online sources, such as waiting time and service rate in the theme park, network topology of transportation network, travel time extracted from Google map, etc. To extend this work in the context of other cities, there should be no difficulties in collecting similar data set. Specifically, in most cases, no extra facilities need to be built for data collection;

2. We have adopted network flow model in this thesis to describe the movement of crowds. This model is flexible to extend to model demand surge in different contexts, such as taxi network, pedestrian network, etc. The advantage of applying network flow model is to quantify the demand surges through flow analysis so that operators can provide optimal strategies accordingly;

3. Proactive and reactive strategies we proposed for managing the surges are practical and easy to be extended. Parameters can be tuned with real-world considerations for different problems. In our experiments, we show the relationships between budget deployed and service quality. From which, operators can assess how much the efforts can be delivered and adjust the deployment of service under a reasonable budget.

1.4 Organization of the Dissertation

The rest of this dissertation is organized as follows. In Chapter 2, I review previous studies that handle demand surges in different aspects. In Chapter 3, I use theme park context to illustrate the first strategy of distributing demands temporally. In Chapter 4, I introduce the second strategy of distributing demands spatially using the public transportation context. In Chapter 5, I introduce the problem of optimizing the portfolio for a mixture of transportation services to handle morning and evening demand surges in a business district. I conclude this dissertation thesis and discuss the future works in Chapter 6.

Chapter 2

Related Work

Demand surges or sudden demand build-ups can easily lead to congestions in the urban city. One of the examples lies in large event situation. An observation [31] shows two clear subsequent waves of large crowds due to people going to and leaving the event venue. Such congestion affects the normal operation of transportation system, for example, in [30], a slow-moving traffic for 10 miles from the concert venue in the UK was caused [31]. Other examples include traffic congestions on expressways during peak hours [12], parking obstructions at parking slot [17], evacuation and egress during the emergency at buildings [43] and so on.

A variety of solutions is proposed to solve surges in different contexts. To manage the congestion of the theme park, Disneyland introduced the Fastpass system, which encourages people to visit popular attractions at designated times by providing express accesses to those who follows the instructions[7]. Such system can offset the waiting time at busy attractions and spread the crowd more evenly. To manage the road traffic, real-time information were provided to facilitate the movement of the large number of passengers [2]. Closure of tunnels and re-route vehicles is another way[2]. Other measures include traffic signal re-timing and real-time traffic monitoring [32].

Strategies to handle demand surges are mainly categorized into two ways: proactive and reactive, with the consideration of their reactions to the situations. Proactive strategy refers to responding to the system before congestions get worse. Reactive strategy is to act to the situation rather than anticipating the future. We summarize some works in Table 2.1

Table 2.1: A summary of approaches in different contexts

Context	**Solution**	**Reference**
Emergency egress	Integrate the desired states and the current prevailing traffic conditions to produce real-time traffic control schemes.	[35]
Theme park	Routing visitors to less crowded areas by offering incentives and information on mobile devices.	[6]
Service systems	Adjust operating capacity (service counters and servers) to for reduce congestion.	[52]
Emergency egress	Maximize the throughput during the specified evacuation duration.	[37]
Traffic congestion	Charge tunnel tolls to adjust the volume of traffic so that reduce congestion.	[61]
Parking congestion	Reduce the parking congestion by increasing the parking fees.	[17]

Among these works, [35], [6] and [52] solve congestion proactively. They measure the crowds and make proper adjustment before the congestion get worse. Whereas [37], [61] and [17] deal with congestion reactively by providing strategies to manage the congestion afterwards.

Although methods are provided, research gaps still remain. Each of them has limitations in various ways and cannot be extended to our domain. For example, [35] and [37] handles surges in emergency egress, whereas in our context, existing infrastructures are in good working conditions and we are trying to complement the existing services. [61] deals with road congestions by altering the charge fees whereas our model mainly focus on public transportation system. Strategy provided in [6] is difficult due to the private issues that visitors' may not want to reveal their current locations. Moreover, adopting this method is costly for park operators as they need to open a developer and data analytics team to handle.

Hence, we study the proactive and reactive strategies to handle demand surges in this thesis. The rest part of this chapter is organized as follows. I review literatures in theme park context in section 2.1 and public transport domain in 2.2.

2.1 Proactive Strategy to Handle Surges in Theme Park Context

2.1.1 Handling Surges in Theme Park

In theme park context, various methods are proposed to handle the congestions. Major ideas include: (1). avoid visitors going to the same attraction simultaneously with the help of a coordination system [26]; (2). balance the capacity at attractions and queuing flow of visitors [1] and so on.

However, the coordination system [26] is based on simulation results. Without real-world information that reflects transitions of visitors, the proposed solutions may not be optimal. The empirical data collected for inferring diffusions is from a survey rather than from operators in [1]. Whereas in our work, visitors' transition behaviours are derived with real world data set offered by park operators, which provides better accuracy for learning the diffusion process.

2.1.2 Dependent Cascade to Model Visitors' Diffusions

I apply Dependent Cascade (DC), which augments Independent Cascade (ID) model, to model visitors transition behaviors. In social network literature, learning the parameters of a IC model based diffusion process has become a flourishing research area [41],[18],[42],[59],[9],[46]. Myers and Leskovec [41] formulate the problem of parameter learning using convex optimization. Gomez *et al.* [18] address a similar problem using sub-modularity based optimization. Netrapalli and Sanghvi [42] address the complementary question of how many observed cascades are necessary to correctly learn the structure of a network. There has also been work on learning the parameters using features of the diffusion process, such as the language of tweets [59] in a Twitter network and geographical features to learn an endangered species movement parameters [60]. Daneshmand *et al.* [9] investigate the network structure inference problem using an l_1-regularized likelihood maximization framework recently. Qu *et al.* [46] utilized data summarization tools to learn the diffusions over large and dynamic online social networks.

Our work is different in the sense that the underlying diffusion process has dependent outgoing flows from a node to maintain flow conservation unlike the IC model. We also assume that only aggregate information is available, rather than individual-level tracking information required in the IC model. The dependent cascade model we use is closely related to the collective diffusion model (CDM) in [29]. Our work addresses an enriched version of such a collective diffusion model as data requirements in our approach are much weaker. In addition, a significant contribution of our work is to formulate and validate diffusion models with real world data, which was not provided in [29].

In the proactive strategy to handle demand surges, with a learned model of the diffusion process, the next step is decision-making in the context of the learned model to handle the congestions proactively The goal is typically to find the set of actions that result in a diffusion process with certain desirable properties subject to operational constraints [27, 28, 50],[36]. This kind of decision-making has numerous applications. For example, in disease control, the decision is which nodes to vaccinate to curb the spread of disease while in advertising, the question is which nodes to target to encourage the adoption of a product. In traffic scenario, the decision is to route cooperate vehicles such that can ease congestions, the challenge then becomes how to capture the dynamic and congestion situation and control the vehicles. Our work proposes and develops an optimization formulation for similar decision making problem within the context of dependent cascades and aggregate flow.

2.2 Reactive Strategy to Handle Surges in Public Transportation Context

2.2.1 Handling Surges in Public Transportation

Our focus for solving the demand surges is mainly on the public transportation field. There are extensive works in the literatures discussing about the strategic planning for public transportation services. The general process consists of three major problems: network design, line planning and timetabling [24]. On solving these three major problems, various approaches and techniques were proposed, readers may refer to the following works for details [33], [5], [25].

Unlike the above works, whose focus were on the strategic planning under normal situations and improving the service quality during a long term period, our work put the emphasis on minimizing the negative effect caused by the demand surges through establishing the temporary services. Reasonable bus planning strategy in context of the normal cases might not be applicable under the demand surge cases as the passengers demand change dramatically during special hours. Moreover, regular bus service that is suitable for the long term period is unnecessary with respect to the special cases, since the impact results from the event only last for a few hours. Therefore, we seek strategic planning for the transportation services under special situations.

One of the special situations in terms of demand surge is metro infrastructure disruption. Contingency plans were investigated in case of disruption, which can be found in [22],[56] and [40]. A survey by Pender *et al.* [44] on the various practices to manage the disruption, which indicated that bus bridging service is the most common way to minimize the negative impact of the disruption. Konstantinos *et al.* [45] proposed a methodological framework for planning the bus services. There were two key steps: bus routes planning on the network and shuttle bus assignment over the selected routes. The optimal bus routes were generated by using the shortest path algorithm and improved by a heuristic approach. Following this framework, Jin *et al.* [24] formulated the problem by applying a different approach for generating candidate bus routes compared to [45].

Despite the similarity between disruption management and regular egress, they are fundamentally different, in the following aspects. For disruption management, the priority is on restoring as much connectivity as possible, and as a result, the modeling effort has been mostly on maximizing the *amount of flow* that can pass through the point of disconnection. For regular egress, on the other hand, the focus is on experience management, which aims at minimizing total journey time including both travel and waiting time.

In this thesis, my major focus is the case where the transportation services are in good working condition. We are trying to provide complementary services rather than contingency plans. However, the proposed methods are general enough to study the demand surges during both normal and train disruption cases. Specifically, the model will generate efficient bus routes for passengers regardless of the completeness of service network.

2.2.2 Studies for Serving Public Transportation

Public transport modes include city buses, trams, massive rapid transit and so on. In this thesis, we studied bus planning and bike sharing planning problems and I will introduce the related studies in following sections.

2.2.2.1 Line planning

As mentioned above, line planning is one of the key procedures in bus service planning problem. Line planning problems mainly discuss the design of routes and their relative frequencies over the public transportation network. In [48], Anita et al. proposed an integer programming approach and apply the Dantzig-Wolfe decomposition method to get the solution. Year after, Ralf et al. presented the branch and price method to solve the problem in [5]. Later in 2012, A comprehensive survey of line planning problem was summarized in [47].

Taking above researchers' work as a basis, I solve the line planning problem in two phases. In the first phase, we make use of column generation technique to help derive good bus candidate routes, which contains a list of candidate bus stops for passengers to take-off. When devising the routes, we also consider the travel delay at each stop. In the second phase, we assign a fleet of buses over each route and distribute passengers in each bus when there are enough capacities allowed. The target is to achieve the minimal travel time of all passengers.

2.2.2.2 Bike Sharing Problem

The bike-sharing system is another transportation option that we addressed in this thesis. It is successfully implemented in many places in the world, such as cities in Europe [3], Asia and America [49]. For implementing the bike-sharing problem, three major aspects are studied: strategic, tactical or operational planning [58]. Strategic planning problems focus on making decisions on the number, location and size of bike stations. Martinez *et al.* [38] optimized the location of biking stations and measured the bike relocation activities required in a regular operation day. Tactical planning takes the variant demand into consideration and try to optimize the service level by maintaining the bike fill level in the station. George *et al.* [16] applied a closed queuing network to model the service network and then find the optimal fleet size and allocation of rental vehicles. On the operational level, relocation issues over shorter terms are studied. In 2013, Dell'Amico *et al.* [10] addressed the Bike sharing Rebalancing Problem (BRP), with a fleet of capacitate vehicles employed to redistribute the bikes.

In our work, the challenge for solving bike-sharing problem not only consists of the above-mentioned issues, moreover, a key problem is to determine the portfolio assignment over both bike and other public transportation services. The way that we connect the two independent planning problem is based on there shared constraints: demand and budget fulfillment constraints. We make use of the Lagrangian Relaxation and Sub-gradient optimization approach to solve them.

Chapter 3

Learning and Controlling Network Diffusion: A Temporal Redistribution Approach

3.1 Overview

3.1.1 Motivation

The tourism and entertainment industry plays an increasingly important role in the global economy. In recent years, theme parks have been an important driver in the growth of this industry. Unfortunately, a vibrant growth in the theme park industry comes hand-in-hand with worsening congestion and increased wait times. From studies, we notice that the wait times on weekends and holidays at popular attractions, particularly in Asian theme parks reaches 2-3 hrs [1].

Improving visitor experience by reducing overall wait times in the theme park context is challenge in two ways. Firstly, a complete mobility model must be learned from aggregate data set. In other words, the park operator records only the waiting time at attractions during different hours of the day, rather than track each patron

[1] http://www.scmp.com/news/china/money-wealth/article/2021968/100-days-crowds-queues-and-some-complaints-after-opening

individually, which, make it hard to understand how people are transiting over different attractions in the park. Secondly, having learned the mobility patterns, we need to provide effective strategy to help reduce the congestions. This strategy must be practical and not too costly.

3.1.2 Problem Description

We provide an operational model to represent diffusion in a time indexed graph, $G(V, E, T)$ with dependent cascades. Before explicitly explaining our model, we start with the well known independent cascade model [27],[50] that is used to represent spread of ideas, influence. Every edge $(u, v) \in E$ at time t is associated with a transition probability $p^t_{u,v}$ representing the probability that node v will be activated if node u was previously activated. In the independent cascade model, probabilities associated with outgoing edges from a node are independent of each other and hence the cascades in different parts of the network are independent.

On the other hand, since we primarily consider diffusion of agents (people/vehicles) where flow is preserved, the probabilities associated with edges going out of a node are dependent on each other. Specifically, we have the following flow preservation dependency for every node u and time t: $\sum_w p^t_{u,w} = 1$. Namely, every agent coming out of u move to one of the nodes w according to the diffusion dynamics, **p**. Our work is concerned with the following learning and planning problems in the context of such dependent cascade models of network diffusion:

(a) Learning: Compute the transition probabilities, **p** that maximize the likelihood of observing the aggregate observations **n** (ex: number of people waiting in queues at attractions in a theme park over multiple days).

(b) Planning: Given **p** and budget B, compute the plan for execution of management actions that achieve a desired objective.

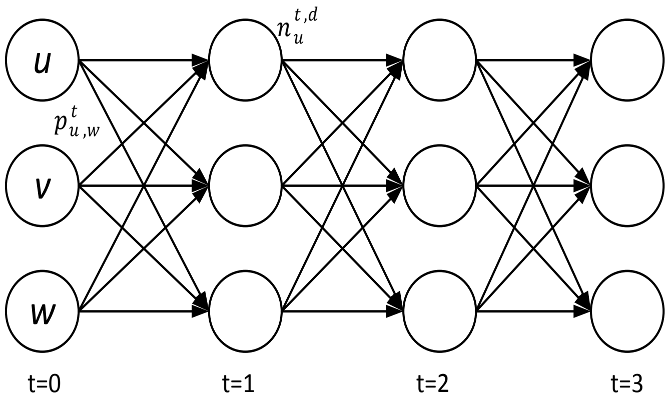


Figure 3.1: Time indexed graph representing diffusion of visitors for 3 time steps.

3.1.3 Application to Theme Park Management

Time indexed graph in figure 3.1 represents diffusion of people at theme park. Each node represents an attraction and an attraction being active at a time step t indicates that there are people who are coming out of the attraction after getting serviced at that time step. $p^t_{u,v}$ associated with an edge between attractions u and v indicates the probability that a visitor coming out of attraction u would move to the attraction v at time t. Since visitors stay within the theme park, summation of transition probability over all outgoing attractions v from u is 1, i.e., $\sum_v p^t_{u,v} = 1$.

At a theme park, it is impractical to track every visitor transitioning between attractions (referred to as $\mathbf{x}$ henceforth). Instead, there are both people and sensors ([2]) to guide and track people at individual attractions. We use this information from multiple days, represented as $n_u^{t,d}$ (number of people waiting at attraction u at time t in day d). The **learning problem** can be formally defined as:

$$\max_{\mathbf{p}} \quad \mathcal{L}(\mathbf{p}|\mathbf{n}) \tag{3.1}$$

where $\mathcal{L}(\mathbf{p}|\mathbf{n})$ represents the likelihood of parameters $\mathbf{p}$ given the observed $\mathbf{n}$.

Theme parks typically conduct moving road shows and/or photo-shoots with cartoon characters near attractions that have high wait times to ensure people spend lesser time waiting at attractions. In doing so, passengers are re-distributed in tem-

[2]http://www.qmetrix.com/sensors/queue-waiting-time-sensors/

poral dimensions and such redistribution helps alter the existing mobility patterns and hence reduce the waiting time. However, current practice of placement of such side shows is ad-hoc and does not consider the diffusion dynamics over time steps. To address this, our planning approach will compute a placement of sideshows, **l**, on edges to minimize wait times while respecting a budget available for side shows. Thus, the **planning problem** can be formally defined as:

$$\min_{\mathbf{l}} \sum_{d,t,u} \omega_u^t(\mathbf{l}, \mathbf{p}) \qquad \textbf{s.t.} \sum_{u,v,k,t} f(l_{u,v}^{t,k}) \leq B$$

where $\omega_u^t(\mathbf{l}, \mathbf{p})$ is the wait time at time step t for attraction u given diffusion dynamics **p** and placement of side shows given by **l**. f represents the cost of placing a side show and B is the budget available.

3.2 Learning Diffusion Dynamics

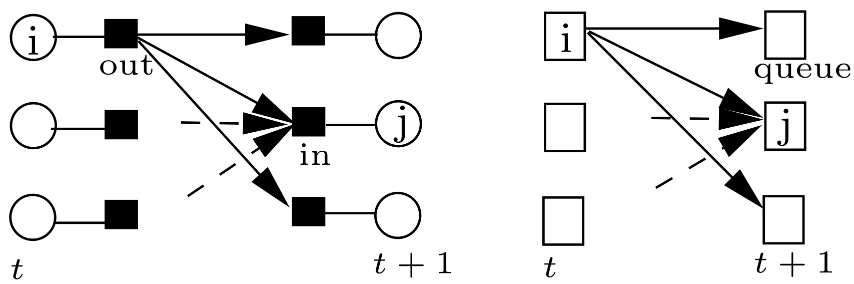
We now describe our method for solving the learning problem described in Equation 3.1, when we only have access to aggregate observations **n**. We assume that the diffusion dynamics from any node u described by $\mathbf{p}_u$ is a standard probability distribution characterized by a few parameters. Specifically, we explore the most relevant ones, namely Multinomial and Poisson. The notation employed in this and subsequent sections is provided in Table 3.1. Boldface letters are used to represent vectors of the items described by the corresponding normal-face letter.

Before we describe our approach, we note that our learning problem is different than the collective flow diffusion (CFD) model of [29] that was developed to understand traffic flow. In the CFD model, it is assumed that the total number of agents that exit and enter each node are known as shown in Fig. 3.2(a). In Fig. 3.2(a) each dark square represents the total in and out flow of agents.

However, in many cases, such as our theme park setting, this type of aggregate data is not available. Instead, we only observe the total number of people waiting to be serviced at an attraction i as provided by the theme park operator. This makes

Table 3.1: Notation

Variable	Definition
$p_{u,v}^t$	Transition probability between nodes u and v at time step t
$n_u^{t,d}$	Number of agents in node u at time t on day d
s_u	Service rate (number of agents serviced in one time step) of node u
D	Set of cascades or days on which observations about **n** are made
$x_{u,v}^{t,d}$	Number of agents moving from node u to node v at time t on day d
$l_{u,v}^{t,k}$	Binary variable that is set to 1 if side show of type k is placed on edge (u,v) at time t
β^k	Percentage of diffusion attracted by side show of type k



(a) (b)

Figure 3.2: a) shows the aggregate data requirement for the CFD model of [29]; b) shows the same for our model.

learning more complex in our case due to more hidden variables than the CFD model. We develop additional constraints to reflect such visitor queues as explained later using constraint 3.5 in Table 3.2. We can further enrich our diffusion models using features of attractions similar to [60]. However, visitors currently do not have access to features such as advanced notification of wait times for attractions they are planning to visit next.

This makes learning more complex due to more hidden variables than the CFD model. We develop additional constraints to reflect such visitor queues as explained later using constraint 3.5 in Table 3.2.

We provide multinomial and poisson distribution based diffusion model and develop optimization based learning algorithm for computing the parameters.

3.2.1 Multinomial Distribution Based Diffusion

Multinomial distribution – a generalization of the binomial distribution – is a categorical distribution where each trial results in exactly one of k possible outcomes with probabilities $P_1, \cdots, P_k$ (so that $P_i \geq 0, \forall i = 1, \cdots, k$) and $\sum_{i=1}^{k} P_i = 1$. We represent diffusion from each node, u at time t as a multinomial distribution with probabilities given by $\{p^t_{u,v}\}_{v \in V}$ for each of the outcomes $v \in V$. Therefore, the probability of observing $x^{t,d}_{u,v}$ number of transitions[3] from node u to node v, $x^{t,d}_{u,w}$ number of transitions from node u to node w and so on for any day/cascade d is given by:

$$Pr(\mathbf{x}^{t,d}_u|\mathbf{p}) = \frac{(\sum_z x^{t,d}_{u,z})!}{\prod_z x^{t,d}_{u,z}!} \prod_z (p^t_{u,z})^{x^{t,d}_{u,z}}$$

where $\sum_z p^t_{u,z} = 1$ and $x^{t,d}_{u,z}$ represents the number of times (frequency) there was a transition from u to z at time t on day d. Since, we do not observe either the probabilities, **p** or frequencies, **x**. We learn them by maximizing the likelihood, $\mathcal{L}(\mathbf{p}|\mathbf{x}, \mathbf{n})$.

More specifically, over all the attractions, likelihood is defined as follows:

$$\mathcal{L}(\mathbf{p}|\mathbf{x}, \mathbf{n}) = Pr(\mathbf{x}, \mathbf{n}|\mathbf{p}) = \prod_{d \in D} \prod_{t \in T} \prod_{u \in V} \frac{(\sum_z x^{t,d}_{u,z})!}{\prod_z x^{t,d}_{u,z}!} \prod_z (p^t_{u,z})^{x^{t,d}_{u,z}} \tag{3.2}$$

where

$$\sum_z x^{t,d}_{u,z} = \begin{cases} n^{t,d}_u & \text{if } \; n^{t,d}_u < s_u \\ s_u & otherwise \end{cases}$$

Given the equivalence of maximizing likelihood and maximizing log likelihood, we employ the following objective

$$\max_{\boldsymbol{p}} \log \sum_{\boldsymbol{x}} Pr(\boldsymbol{n}, \boldsymbol{x}|\boldsymbol{p}) \tag{3.3}$$

[3]Since **x** represents an observation, it can be different for different days or cascades, $d \in D$.

Table 3.2: GETDIFFUSIONDYNAMICS($\boldsymbol{n}$, $\boldsymbol{s}$)

$$\textbf{max:} \sum_d \sum_u \sum_t \left(\log\left((\sum_z x_{u,z}^{t,d})!\right) - \sum_z \log(x_{u,z}^{t,d}!) + \sum_z x_{u,z}^{t,d} \log(p_{u,z}^t) \right)$$

$$\textbf{s.t. } n_u^{t+1,d} = n_i^{t,d} + \sum_z x_{z,u}^{t,d} - \sum_z x_{u,z}^{t,d}, \quad \forall t, d, u \tag{3.5}$$

$$\sum_z x_{u,z}^{t,d} \leq min(s_u, n_u^{t,d}), \;\; \forall t, d, u \tag{3.6}$$

$$\sum_z p_{u,z}^t = 1, \;\; \forall t, u \tag{3.7}$$

$$x_{u,z}^{t,d} \in \mathbb{N}_0, \;\; \forall t, d, u, z \tag{3.8}$$

$$0 \leq p_{u,z}^t \leq 1, \;\; \forall t, u, z \tag{3.9}$$

To make it computationally simpler, we use the following approximation:

$$\max_{\boldsymbol{p},\boldsymbol{x}} \log Pr(\boldsymbol{n}, \boldsymbol{x}|\boldsymbol{p}) \tag{3.4}$$

This approximation is in the same spirit as the one in Sheldon *et al.* [51]. The main intuition is that for categorical distributions, such as Binomial distribution, the mode is very close to the mean. The above optimization problem can be formulated as a non-linear program as shown in Table 3.2. The objective function is the logarithm (log) of Eq. (3.2). The first and the second constraint jointly represent the flow conservation at each node. In the first constraint, the number of visitors at a node u at time $t+1$ according to a cascade d (= $n_u^{t+1,d}$) is constrained to be equal to the number of visitors at the same node at time t with the addition of in-flow into the node (=$\sum_z x_{z,u}^{t,d}$) and subtraction of the out-flow from the node (=$\sum_z x_{u,z}^{t,d}$) at the same time step. The second constraint ensures that the out-flow is equal to the minimum of the service rate at node u and the number of people currently waiting to be serviced at the node at time step t. Rest of the constraints enforce basic properties of the diffusion model.

We solve the optimization problem in Table 3.2 using a commercial non-linear solver called Lingo (http://www.lindo.com).

3.2.2 Poisson Distribution Based Diffusion

Poisson distribution is often used to provide the probability of a given number of events occurring in a fixed time interval. Therefore, we also model the visitor diffusion process using this distribution. Similar to the previous two models, we first define the log-likelihood as follows:

$$\mathcal{L}(\lambda|\boldsymbol{x},\boldsymbol{n}) = P(\boldsymbol{x},\boldsymbol{n};\lambda) = \prod_{d\in D}\prod_{t\in T}\prod_{i\in A}\prod_{j\in A}\frac{\lambda_{t,i,j}^{x_{d,t,i,j}}e^{-\lambda_{t,i,j}}}{x_{d,t,i,j}!} \tag{3.10}$$

Taking log of the above, we get:

$$\log(\mathcal{L}(\lambda|\boldsymbol{x},\boldsymbol{n})) = \sum_{d,t,i,j}\Big[x_{d,t,i,j}\log(\lambda_{t,i,j}) - \lambda_{t,i,j} - \log(x_{d,t,i,j}!)\Big]$$

The key parameter of interest in addition to **x** with Poisson distribution is $\boldsymbol{\lambda}$. The optimization formulation in Table 3.2 is appropriately modified to reflect this change.

3.3 Controlling Diffusion Dynamics Through Redistributing Visitors in Temporal Dimension

We now describe our mechanism to compute plans of management actions that will be used to control diffusion dynamics through re-distributing the visitors. In this work, we consider management actions that can be viewed as dampeners that are placed on an edge to absorb the diffusion on that edge for a certain time duration. In the context of a theme park, these management actions correspond to side shows that can be placed between attractions for a limited time. Depending on their type, such management actions have an associated cost and impact on the diffusion. For instance, a photo opportunity with a cartoon character only attracts a few people and is not typically expensive as only one actor is involved. In contrast, an elaborate road show attracts most visitors traveling on that edge, and is more expensive due to multiple actors being involved.

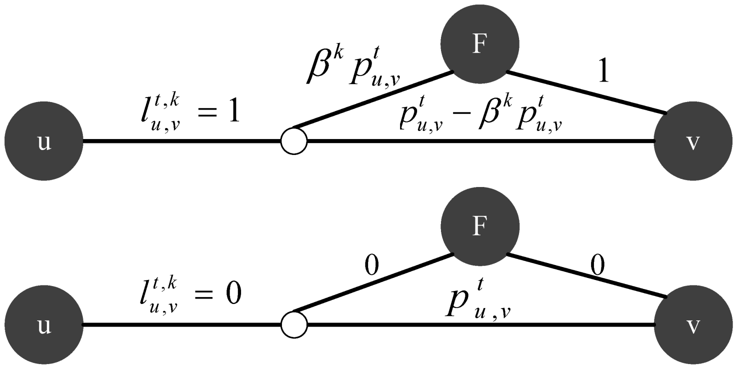


Figure 3.3: Representation of management actions

Figure 3.3 provides a visual representation of how the diffusion dynamics are altered due to a management action. As explained earlier, $p^t_{u,v}$ represents the preference of visitors coming out of node u to move to node v; β^k is the proportion of visitors absorbed on an edge for one time step due to the management action of type k. When $l^{t,k}_{u,v} = 1$, there is a management action of type k executed on edge between u and v at time step t. In this case, β^k proportion of visitors are absorbed into the buffer node F for the current time step t. Their movement is re-distributed in the temporal dimension. After one time step, at time step $t + 1$, such visitors at the buffer node move with probability 1 to their original intended destination v. Thus, introducing the buffer node via the action $l^{t,k}_{u,v} = 1$ helped alter the mobility patterns and hence reduced congestion at node v at time step t. Action $l^{t,k}_{u,v} = 0$ indicates that there is no management action of type k executed on edge between u and v at time step t. In this case, the diffusion probability remains the same as $p^t_{u,v}$.

Given the diffusion dynamics, **p**, the goal is to compute the plan **l** of management actions that will minimize the total waiting time or latency over all the nodes across all time steps over different realizations of diffusion. We also assume that each management action has a cost, and there is a fixed available budget. Two existing methods that have been employed for controlling/influencing diffusion are:

- Exploiting sub-modularity: (*sub-modularity is a property of set functions indicating that the difference in the incremental value of the function that a single element makes decreases as the size of the input set increases.*) The problem of selecting a fixed number of nodes in a social network that will maximize influence in the context of independent cascade model [27] is solved by exploiting sub-modularity of the objective. Specifically, because of sub-modularity, nodes can be greedily selected one after another based on their marginal addition to the overall influence. Such an approach provides a solution that is at least $1 - \frac{1}{e}$ (=63%) of the optimal. Unfortunately, with dependent diffusion dynamics, the problem of minimizing wait time is not sub-modular. (I will show the proofs in the Appendix.)

Given the relevance, we experimentally benchmark the performance of our approach against the greedy approach, even though it does not provide quality guarantees.

- Employing Sample Average Approximation (SAA): The problem of buying parcels of land to maximize the population of rare species [50] is formulated as a stochastic optimization problem. The key idea is that instead of solving the stochastic optimization problem directly, a solution is computed for a few samples from the diffusion process. Because of independence in cascades, samples can be generated before the optimization.

  Unfortunately, when simulating a large population of visitors moving in a theme park, we would need to sample the trajectory for each of them in the context of SAA. This leads to a prohibitively large size of the decision problem when formulated using a mathematical program, and is not scalable.

Therefore, the main contribution of this work with respect to controlling diffusion dynamics is a scalable approach that substitutes the computation of expected wait time (expectation over trajectory samples of visitors) with wait time for expected numbers of visitors (expectation over diffusion dynamics) over all nodes and all time steps. Specifically, we denote $E_{\boldsymbol{P}}[\sum_{i,t} g(n_i^t)]$ as the expected wait time for the joint multinomial distribution $\boldsymbol{P}(= \prod_t \prod_i \frac{(\sum_j x_{i,j}^t)!}{\prod_j x_{i,j}^t!} \prod_j (p_{i,j}^t)^{x_{i,j}^t})$ over all nodes and all time steps. We use $\sum_{i,t} g(E_{q_i^t}[n_i^t])$ to represent the wait time of expected number of visitors with distribution q_i^t for each individual attraction i at time t.

Proposition 3.3.1 *With liner function $g(x)$ measuring the wait time (i.e. $g(x) =$*

$ax + b$), $E_{\boldsymbol{P}}[\sum\limits_{i,t} g(n_i^t)] = \sum\limits_{i,t} g(E_{q_i^t}[n_i^t])$.

Proof First of all, we explain the equivalence of $E_{\boldsymbol{P}}[\sum\limits_{i,t} g(n_i^t)]$ and $\sum\limits_{i,t} g(E_{q_i^t}[n_i^t])$:

$$\begin{aligned} E_{\boldsymbol{P}}[\sum_{i,t} g(n_i^t)] &= \sum_{i,t} E_{q_i^t}[g(n_i^t)] \\ &= \sum_{i,t} E_{q_i^t}[a * n_i^t + b] \\ &= \sum_{i,t} a * E_{q_i^t}[n_i^t] + b \\ &= \sum_{i,t} g(E_{q_i^t}[n_i^t]) \end{aligned}$$

The expected wait time can be easily obtained if we get $E_{q_i^t}[n_i^t]$. ■

To simplify the notation, we use $E[n_i^t]$ instead of $E_{q_i^t}[n_i^t]$. We propose the following proposition.

Proposition 3.3.2 *If* $n_i^t \geq s_i$ *for all* $1 \leq i \leq m$ *and* $t \geq 1$, *we have*

$$E[n_i^{t+1}] = E[n_i^t] + \sum_{j=1}^{j=m} \left(p_{j,i} \cdot s_j\right) - s_i \tag{3.11}$$

Proof Obviously, we have $E[n_i^{t+1}] = E[n_i^t + \sum_{j=1}^{j=m}(x_{j,i}^t - x_{i,j}^t)] = E[n_i^t] + \sum_{j=1}^{j=m}\left(E[x_{j,i}^t] - E[x_{i,j}^t]\right)$, where $x_{i,j}^t$ is the number of people move to attraction j after they finished the attraction i at time t. So the problem becomes how to calculate $x_{i,j}^t$.

$$
\begin{aligned}
E[x_{i,j}^t] &= \sum_{k=0}^{k=\infty}\sum_{y=0}^{y=k}\big(P(n_i^t = k, x_{i,j}^t = y|s_i)\cdot y\big) && (3.12)\\
&= \sum_{k=0}^{k=\infty}\sum_{y=0}^{y=k}\big(P(n_i^t = k)\cdot P(x_{i,j}^t = y|n_i^t = k, s_i)\cdot y\big) && (3.13)\\
&= \sum_{k=0}^{k=\infty}\big(P(n_i^t = k)\cdot \sum_{y=0}^{y=k} P(x_{i,j}^t = y|n_i^t = k, s_i)\cdot y\big) && (3.14)\\
&= \sum_{k=0}^{k=\infty}\big(P(n_i^t = k)\cdot \sum_{y=0}^{y=\gamma} \frac{\gamma!}{y!(\gamma-y)!} p_{i,j}^y(1-p_{i,j})^{\gamma-y}\cdot y\big) && (3.15)\\
&= \sum_{k=0}^{k=\infty}\big(P(n_i^t = k)\cdot \gamma\cdot p_{i,j}\big) && (3.16)\\
&= p_{i,j}\cdot\sum_{k=0}^{k=\infty}\big(P(n_i^t = k)\cdot \gamma)\big). && (3.17)
\end{aligned}
$$

where $\gamma = \min(k, S_i)$ in Eq. (3.15). The first "=" is from the independent relationship among the joint multinomial distributions over a set of attractions.

Since $n_i^t \geq S_i$ by assumption, we can easily get $P(n_i^t = k) = 0$ when $k < S_i$. Thus, $E[x_{i,j}^t] = p_{i,j}\cdot\sum_{k=S_i}^{k=\infty}\big(P(n_i^t = k)\cdot\gamma)\big) = p_{i,j}\cdot\sum_{k=S_i}^{k=\infty}\big(P(n_i^t = k)\cdot S_i)\big) = p_{i,j}\cdot S_i$. Therefore,

$$
\sum_{i=0}^{i=m} E[x_{i,j}^t] = \sum_{i=1}^{i=m} p_{i,j}\cdot s_i, \qquad (3.18)
$$

$$
\text{and} \quad \sum_{i=0}^{i=m} E[x_{j,i}^t] = \sum_{i=1}^{i=m} p_{j,i}\cdot s_j = s_j. \qquad (3.19)
$$

Therefore, the Proposition 3.3.2 holds.

Further, when $t = 1$, we have $E[n_i^{t=1}] = n_i^{t=1}$, for all $1 \leq i \leq m$. Thereby, we can recursively obtain $E[n_i^t]$ based on Eq. (3.11), for all $1 \leq i \leq m$ and $t \geq 2$. ■

The resulting optimization formulation is much smaller and scalable when compared with the sampling approach using SAA. It should however be noted that even with using expected numbers of visitors, the problem of minimizing wait time using management actions remains NP-Hard.

Proposition 3.3.3 *The problem of minimizing wait time for expected numbers of visitors at nodes over all time steps by using management actions **l** and a given budget is an NP-Hard problem.*

We omit the proof. We prove this by showing that 0/1 knapsack problem is a special case of our problem. Specifically, we reduce the 0/1 knapsack problem to our problem. The key insight is that minimizing wait time at all nodes at all times is equivalent to maximizing number of agents in buffer nodes at all times. By mapping items in knapsack to management actions, weight of an item, w^k to the cost, c^k of executing management action and value of an item and finally v^k to the overall increase in number of agents at buffer node due to execution of management action, we demonstrate this reduction. ■

Table 3.3 provides an optimization formulation to compute **l** that minimizes the average wait time for expected number of visitors at every node and at every time step. Let $|U|$ denote the total number of attractions in the theme park, and T denote total time steps. Each time step in our case denoted a block of 1 hour period during the day resulting in $T = 9$. The objective function is the average wait time over all the attractions u across all time steps. This particular metric was suggested to us by the theme park operator and is a key performance indicator for the theme park management. While this formulation contains non-linear constraints, we will subsequently provide linear equivalents. We refer to this approach as CDON (Controlling Diffusion through OptimizatioN).

Constraints (3.20) ensures that expected number of visitors in a node, u at time $t+1$ is equivalent to the expected number of visitors at u at time t plus the expected number of visitors transitioning out of u minus the expected number of visitors transitioning into u at time t. Constraints (3.21) are the flow preservation constraints, which ensure that all visitors coming out of a node go to one of the other nodes. This introduces dependencies across cascades. Constraints (3.22) ensures correct computation of number of visitors coming out of a node u. Service rate s_u indicates the number of visitors that are served in one time step. Therefore, the number of visitors

Table 3.3: CDON($\boldsymbol{p}$, $\boldsymbol{s}$, $\boldsymbol{n^0}$)

$$\textbf{min:} \frac{1}{|U| \cdot T} \sum_{u,t} \frac{n_u^t}{s_u}$$

$$\textbf{s.t.} \quad n_u^t + \sum_z x_{z,u}^t - \sum_z x_{u,z}^t = n_u^{t+1} \qquad \forall u, t \tag{3.20}$$

$$\sum_z x_{u,z}^t = y_u^t \qquad \forall u, t \tag{3.21}$$

$$y_u^t = min(s_u, n_u^t) \qquad \forall u, t \tag{3.22}$$

$$x_{u,z}^t = y_u^t \cdot \left[p_{u,z}^t - \sum_k \beta^k \cdot l_{u,z}^{t,k} \cdot p_{u,z}^t \right] \qquad \forall u, z, t \tag{3.23}$$

$$x_{u,F}^t = y_u^t \cdot \left[\sum_k \beta^k \cdot l_{u,z}^{t,k} \cdot p_{u,z}^t \right] \qquad \forall u, z, t \tag{3.24}$$

$$x_{F,z}^t = \sum_u \sum_k \beta^k \cdot l_{u,z}^{t-1,k} \cdot y_u^{t-1} \cdot p_{u,z}^{t-1} \qquad \forall t, z \tag{3.25}$$

$$\sum_k l_{u,z}^{t,k} \leq 1 \qquad \forall u, z, t \tag{3.26}$$

$$\sum_{t,u,z,k} l_{u,z}^{t,k} c^k \leq B \tag{3.27}$$

$$l_{u,u}^{t,k} \in \{0, 1\} \qquad \forall u, v \neq u, k, t \tag{3.28}$$

served in any one time step is the minimum of s_u and number of visitors in u at time t, i.e., n_u^t. Constraints (3.23)-(3.25) clearly state the redistribution procedure and they ensure correct computation of the expected number of visitors[4] that move to other nodes and the buffer node due to placement of side shows. Constraints (3.26) and (3.27) are the constraint on for management actions. Constraints (3.26) ensures that for the same link at the same time, it is only allowed for one type of action. Constraint (3.27) enforces the budget constraints for the management actions.

As can be noted, Constraints (3.22),(3.23),(3.24),(3.25) all contain non-linear terms. The first non linear term is in Constraints (3.22). To provide a linearisation

[4]When there are no side shows, expected number of visitors moving from node u to z at time t is given by $y_u^t \cdot p_{u,z}^t$

to this constraint, we use two binary variables, namely, d_u^t and e_u^t as follows:

$$\begin{aligned} & d_u^t + e_u^t = 1; \quad y_u^t \leq n_u^t; \quad\quad y_u^t \leq s_u \\ & y_u^t \geq n_u^t - M \cdot (1 - d_u^t); \\ & y_u^t \geq s_u - M \cdot (1 - e_u^t) \end{aligned} \tag{3.29}$$

In these constraints M is a large positive number and the tightest bound for M is the largest value of n_u^t and s_u. The validity of these linear constraints can be ascertained by considering all possible values for the two binary variables, d_u^t and e_u^t. Next, we consider the non-linearity in constraints (3.23),(3.24) and (3.25), which is the term $y_u^t \cdot l_{u,z}^{t,k}$. Since y_u^t is positive and $l_{u,z}^{t,k}$ is a binary number, the linear equivalent constraints are:

$$\begin{aligned} & r_{u,z}^{t,k} \leq l_{u,z}^{t,k} \cdot M; \quad r_{u,z}^{t,k} \leq y_u^t; \\ & r_{u,z}^{t,k} \geq y_u^t - (1 - l_{u,z}^{t,k}) \cdot M; \quad r_{u,z}^{t,k} \geq 0 \end{aligned}$$

Again the validity of these constraints can be ascertained by considering both possible values for $l_{u,z}^{t,k}$.

3.4 Experiments: Learning Diffusion Dynamics

In order to demonstrate the utility of our approaches in computing diffusion dynamics, we use a 5-month long data set of wait times from a real theme park in Singapore which consists of 9 major attractions. Using the wait time data and service rate for each attraction in the data set, we get an estimate of how many visitors are currently waiting in queue at each attraction. We are unable to provide a map of the attractions due to confidentiality agreements.

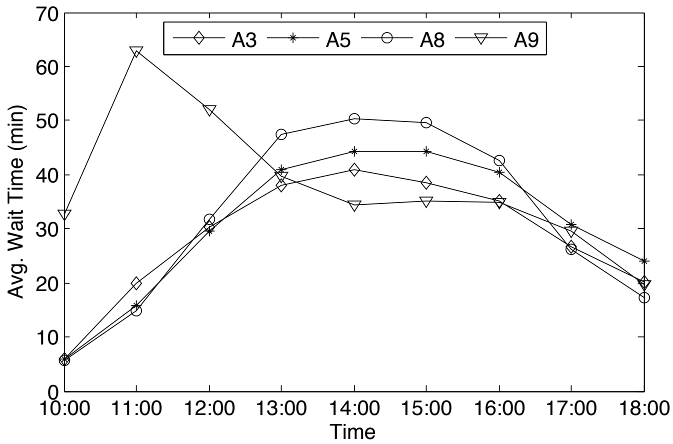


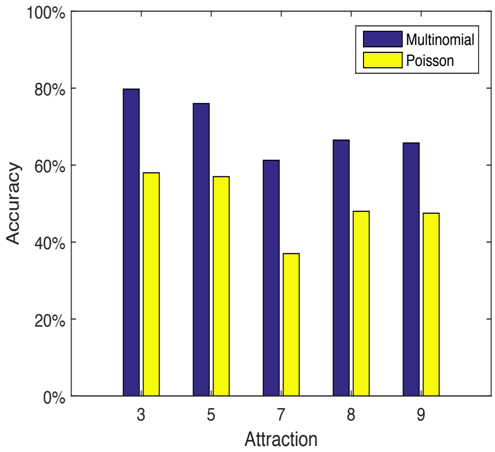
Figure 3.4: Hourly wait time data for the four busiest attractions.

3.4.1 Wait Time Data

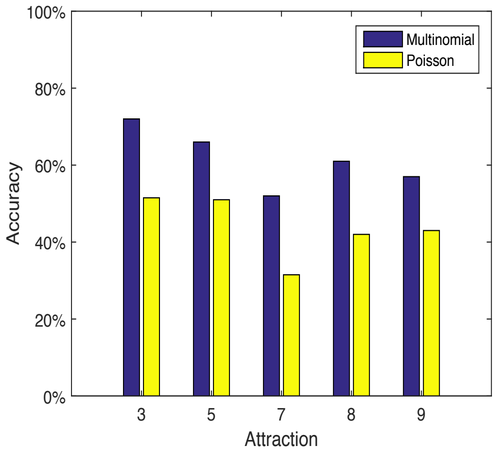
We have wait time data available for each of the 9 attractions. The intervals at which wait times are stored for each of the attractions is different. Therefore, to ensure consistency, we consider the wait time at hourly interval for each attractions. A smaller reporting interval may not be necessarily helpful given that the wait time during the busy period can be quite high.

Fig. 3.4 shows the average wait time for the 4 busiest attractions for each hourly interval starting from 10AM to 6PM. We can clearly see that wait times are close of 1 hour for many of these attractions. Wait times gradually peak during the afternoon hours when the rush is maximum. Attraction 9 ('A9') is the most popular attraction with patrons preferring to visit it early in the day. This causes an early wait time peak for this attraction. Such high wait times motivate our approach to find the mobility patterns of visitors, which can then be analyzed to better manage congestion. Some congestion control measures include conducting sideshows between attractions that have high visitor flow among themselves as well as high wait time.

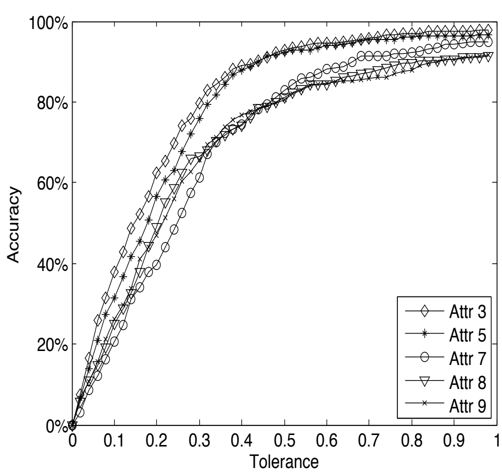
Leisure node: To account for lack of data on visitors entering and exiting the theme park as well as taking breaks, we introduce a new attraction called the 'leisure' node numbered 'A10'. This node is required to account for initial inflow of visitors and their exit. This node has infinite service rate and infinite capacity. We will show



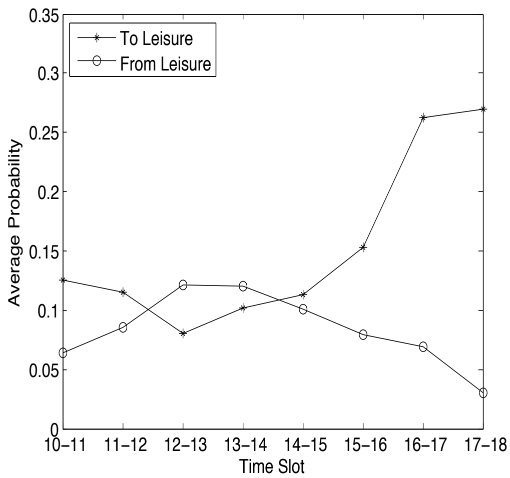
(a) 25% tolerance



(b) 30% tolerance



(c) REC curve



(d) Overview of leisure node

Figure 3.5: Accuracy for 5 busiest attractions and leisure node

concretely how this can be captured using the leisure node.

Accuracy Figures 3.5(a) and (b) show the average accuracy achieved by different diffusion models using the 5-fold cross validation. Using the learned model parameters, for example, transition probabilities $\{p^t_{u,z}\}$ for the multinomial diffusion, we predict the number of people $\boldsymbol{n}$ waiting at each attraction at hourly intervals for all the days in the test data.

To compute the accuracy, we consider a fixed confidence interval. A predicted aggregate value n is considered correct for an attraction if it is within a particular threshold, say 25%, of the true n. Using this definition, we count the total accuracy for all the predictions with one prediction for each time step for each test day per attraction. Figure 3.5(a) and (b) show the accuracy for 25% and 30% threshold. We make the following observations.

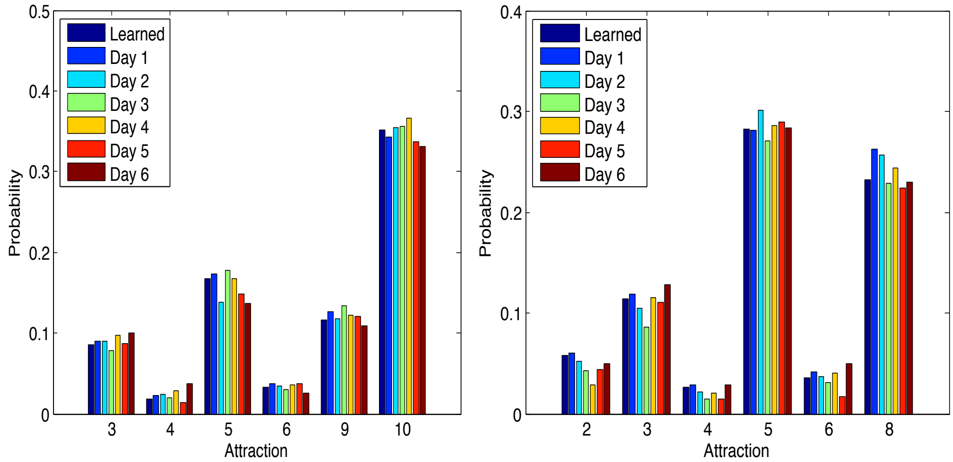
A key observation from figures 3.5(a) and (b) is that attractions that have high wait times (3, 5, 8, and 9 in Fig. 3.4) also have a high accuracy of prediction ($\approx$ 70%-80%). The accuracy is lower for attractions that are relatively lightly congested, such as attraction 7. Attraction 7 as well as attractions 1, 4 and 6 have an average wait time of 15 minutes. This may contribute to lower accuracy as we use hourly wait times. Using finer grained reporting intervals (such as every 15 min.), we expect to increase the accuracy for such lightly congested attractions. We are currently in the process of collecting such fine grained data from the theme park operator. Importantly, our approach is able to provide good accuracy for all the heavily congested attractions validating our models and learning algorithms.

We also observe from figures 3.5(a) and (b) that the multinomial distribution consistently provides higher accuracy than Poisson distributions. The probability mass function for Poisson distribution is as follows:

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!} \tag{3.30}$$

where x is the number of times an event occurs in an interval and x takes integer values. According to the definition of poisson distribution, one of the key characteristics is two events cannot occur at exactly the same instant. In our model, event x represent the action that a visitor coming out of an attraction. However, when finishing the entertainment, many of the visitors may come out of the attraction simultaneously, which implies that the poisson distribution may not proper to model the diffusions in our context. Thus, the results are worse.

Figure 3.5(c) provides the regression error characteristic (REC) curve [4] for the 5 busiest attractions to show how the accuracy varies continuously with the tolerance threshold. The x-axis denotes the tolerance, y-axis denotes the fraction of total predictions that are correctly predicted to be within the tolerance threshold. This graph shows that the accuracy increases sharply between 20%-30% tolerance level.



(a) Parameters for A8 for 4PM-5PM (b) Parameters for A9 for 12PM-1PM

Figure 3.6: Learned parameter verification for the multinomial based diffusion (figures best viewed in color)

As expected, all the busiest attractions (3, 5, 8 and 9) have high accuracy. Lightly congested attraction 7 has lower accuracy. Overall, the above set of experiments indicate a strong performance by our models and algorithms. Using finer grained wait times from the theme park operator, we expect that our approach would provide good accuracy even for lightly congested attractions.

Learned Parameter Verification We now explain the transitions into the leisure node and provide a verification mechanism for the learned parameters obtained for the multinomial diffusion model ($\{p^t_{u,z}\}$). Figure 3.5(d) shows the average transition probability *from* and *to* the leisure node over all the attractions for each time interval of the day. The legend 10-11 refers to the interval 10AM to 11AM. This figure clearly shows that during the beginning of the day, the transitions from leisure node to major attractions gradually increases until reaches the peak at around 12PM to 1PM. Whereas the transition probability to the leisure node is quite low. This is expected as during the morning and early afternoon, visitors arrive in to the theme park and they are joining the queue of those major attractions. In contrast, the transition probability to the leisure node increases significantly towards the latter part of the day. This is also expected as visitors exit the theme park during late afternoon and evenings. Thus, the concept of leisure node is able to capture such visitor movements succinctly.

To verify the parameters generated using the optimization formulation in table 3.2, we compare the learned parameters $\boldsymbol{p}$ against the $\boldsymbol{p'}$ calculated for each day from the $\boldsymbol{x}$ values for that specific day: $p'^{t,d}_{u,z} = \frac{x^{t,d}_{u,z}}{\sum_z x^{t,d}_{u,z}}$. Ideally, the learned parameters $\boldsymbol{p}$ and parameters $\boldsymbol{p'}$ for each test day should be as close as possible.

Figure 3.6(a) and (b) show these comparisons for attraction 8 for time interval 4PM-5PM and attraction 9 for time interval 12PM-1PM respectively. The x-axis denotes the attractions to which visitors can transition to. For example, for figure 3.6(a), $x = 3$ implies the parameter $p'^{t=7,\cdot}_{u=8,z=3}$, where the holder '$\cdot$' is for day number. Intuitively, this parameter represents the probability that a visitor currently at attraction 8 moves to attraction 3 during the time interval 7 (4PM-5PM). For each cluster on the x-axis, we show 7 bars. The first bar corresponds to the 'Learned' parameter from table 3.2. Other bars show the computed parameter p' for different test days, 6 in total.

We make the following observations from figures 3.6 (a) and (b). First, both the learned parameters $\boldsymbol{p}$ and the computed parameters $\boldsymbol{p'}$ are very close to each other. This is true for other busy attractions as well as rest of the time intervals. This further validates our approach. In addition, for figures 3.6(a), we see clearly that transition to the leisure node (x=10) dominates all the other transitions. This is as expected as during the evening hours, visitors exit the theme park. We also see a clear domination of a few attractions to which the visitors move from both the attractions 8 and 9. For example, figure 3.6(b) shows that visitors prefer to move to attractions 3, 5 and 8. Thus, our approach is able to extract meaningful visitor dynamics from the observed aggregate data.

3.5 Experiments: Controlling Diffusion Dynamics

We now provide an empirical evaluation to demonstrate the utility of our approach presented in Table 3.3 to control diffusion with temporal redistribution. This method

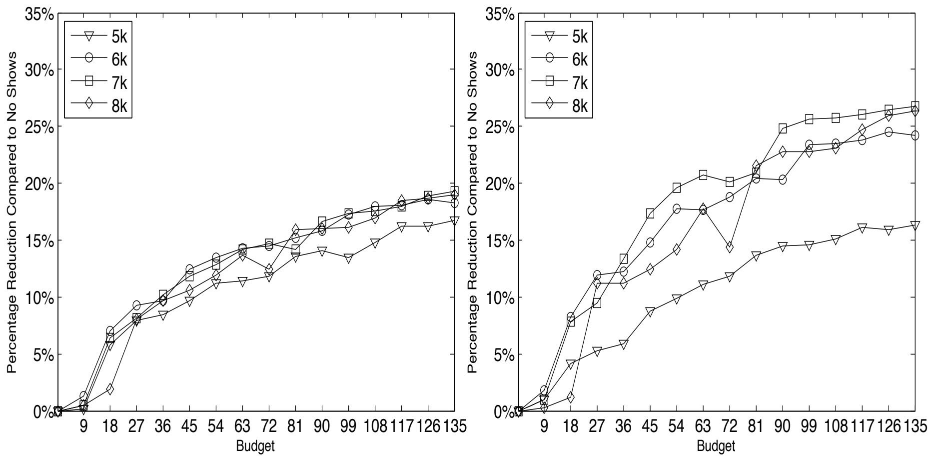


Figure 3.7: Average and peak wait time reduction for CDON

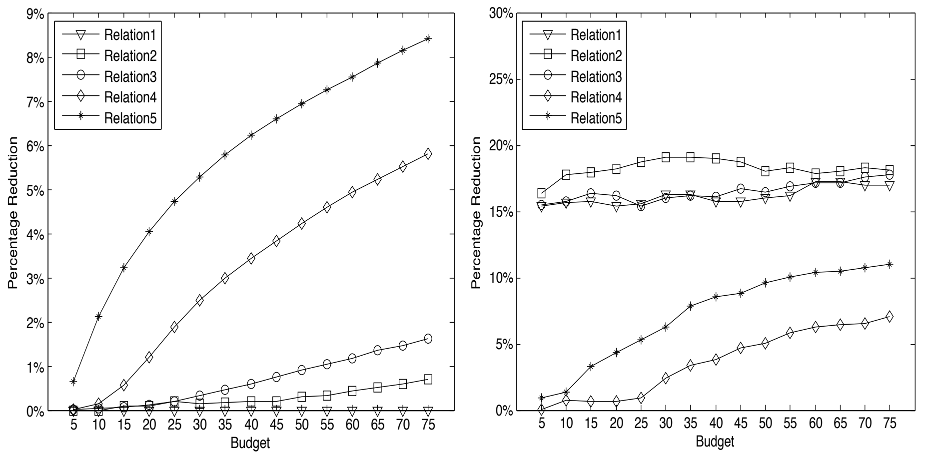
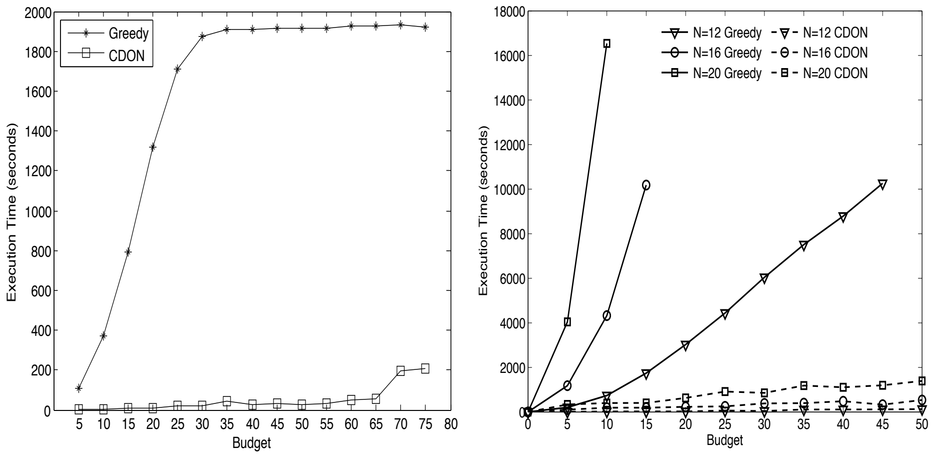


Figure 3.8: Average (left figure) and peak (right figure) wait time reduction due to CDON in comparison with greedy on real data

is compared to the setting of no side shows and the greedy baseline described below. We first consider the real data set of theme park along with the diffusion dynamics obtained using our learning mechanism (described in previous section) and then consider synthetic problems to demonstrate scalability of our approach in comparison to the greedy baseline.

Greedy We employ a greedy algorithm as a baseline in our experiment. It starts with an empty action set $\mathcal{S}$ and iteratively adds the *best* management actions until cost of management actions exceeds the budget. In each round, the objective is evaluated by simulating the cascade when adding every possible management action of type k to $\mathcal{S}$. The action with the largest reduction in wait time is added to the set $\mathcal{S}$.



(a) Real Dataset (b) Synthetic Dataset

Figure 3.9: Runtime comparison of CDON with greedy on real and synthetic datasets

3.5.1 Real-world Dataset

Since side shows are placed on the edges, the complexity of controlling diffusion is dependent on the number of edges, which is 729 in the time indexed graph for the real data set.

Single Type of Side Show Initially, we show the utility of using side shows by considering only one type of side shows and $\beta = 1.0$ for that side show type. Since there is only one type, we consider the budget to be the maximum number of available side shows that can be employed across different time steps. Figure 3.7 shows the average[5] and peak wait time[6] reduction when compared to wait times without any side shows for total population ranging from 5k-8k. As expected, both average and peak wait time reduction increased with the budget, irrespective of the population size and this reduction in wait time is as much as 20% for the average and up to 25% for the peak wait time.

Multiple Type of Side Shows We now consider multiple types of side shows, where the cost, c^k and diffusion absorption parameter, β^k for the type k are connected to each other according to one of the five relations below: (1) $c^k = 10\sqrt{\beta^k}$; (2)

[5] Average wait time is the objective of CDON

[6] Peak wait time is the current wait time with CDON strategy for the attraction that had the highest wait time previously.

$c^k = 10\beta^k$; (3) $c^k = 10(\beta^k)^2$; (4) $c^k = 10(\beta^k)^3$; and (5) $c^k = 10(\beta^k)^4$. Their relations are based on the fact that a side show with low attractiveness (or low β) should be cheaper to deploy than a popular one. We consider $k = 4$ different types of shows are available, with β^k values given by $\{0.2, 0.5, 0.75, 1.0\}$ and a fixed initial population of 7000.

In figure 3.8, we compare the wait time reduction due to CDON in comparison with greedy approach on real dataset or $(W_{greedy} - W_{cdon}) * 100/W_{greedy}$, W_{alg} denotes the objective for the corresponding approach. A higher value of this percentage reduction denotes better performance by CDON over greedy. As the budget is increased, CDON performed consistently better than greedy on all five relations. We observe that as power of β is increased in the relation, the average wait time reduction is increased. Because CDON coordinates the placement of shows at different edges, if we have more available shows to place, the difference with greedy increases. With higher powers for beta, cost values for side shows are smaller resulting in more available side shows. This explains the reason for up to 9% wait time reduction provided by CDON in comparison with greedy. We similarly compare the wait time reduction for the peak attraction due to CDON in comparison with greedy approach in Figure 3.8(b). We observe that for all relations, the wait time reduction percentage is positive and in the best case it goes up to 20%.

We also record the runtime for the 5 different relationships for both CDON and greedy approach, the results are shown in figure 3.9. Due to space constraints, we only provide the results for relation 4. In each relation, greedy takes significantly more time to run than our approach, with the difference increasing as power of β is increased. This is because greedy strategy needs to evaluate many options as more number of side shows can be placed within the same budget.

3.5.2 Synthetic Data Set

To further demonstrate the performance improvement provided by CDON in comparison with greedy, we generate synthetic problems. Our goal is to identify scal-

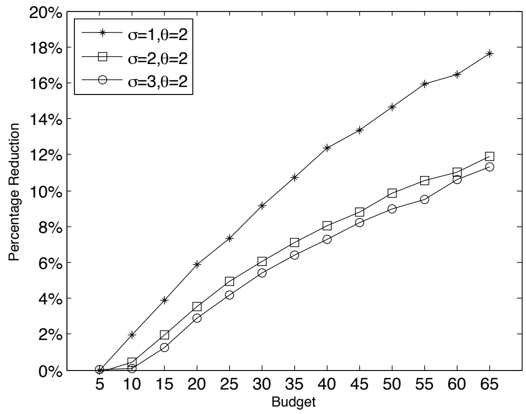


Figure 3.10: Synthetic dataset: reduction in wait time with CDON in comparison to the greedy approach

ability limits of CDON and greedy. The diffusion dynamics are skewed and are generated by using a gamma distribution (with different values of shape parameter σ and scale parameter θ). The sum of diffusion probabilities out of an attraction are normalised to 1. We show results for relation 4.

Figure 3.9(b) provides the runtime results as the number of nodes (N=12,16,20) and budget are increased. We have runtime on y-axis and budget on x-axis. There is an order of magnitude reduction in runtime provided by CDON in comparison with greedy. For most cases, greedy does not compute a solution within our threshold of 10000 seconds. This is because greedy has to evaluate placement of a side show of each type on each of the edges in the time indexed graph, and the number of edges has been increased from 729 (real problem) to 1296 (for N=12), 2304 (for N=16) and 3600 (for N=20). Thus, our CDON approach is highly scalable w.r.t. the number of attractions or network nodes as opposed to greedy. Figure 3.10 provides the percentage reduction in average wait time as the budget is increased for the same relation with different values of the gamma distribution's parameters. We again observe that gain by CDON increases up to 15% as the budget is increased.

3.6 Conclusion and Discussion

Managing diffusion in networks is an important and challenging problem with applications in ecology, leisure and entertainment, and marketing among others. Existing work has primarily focused on phenomena that diffuse independently on all outgoing edges of a node. We augmented the basic independent cascade model with important features required to model real-world problems, such as learning from aggregate data, modeling queues at network nodes and addressing flow conservation at network nodes. We also developed an optimization based approach that provided a plan for management actions to distribute the underlying visitors' diffusion process in a theme park for reducing the average wait time. We also demonstrated the efficiency and effectiveness of our learning and planning approaches through extensive evaluations on both real world and synthetic problems.

A limitation of using HMM to model diffusions is: future transitions only depend on the current state and are independent of the history. To account for such limitation, Long Short-Term Memory (LSTM) [20] model can be utilized. Typically, LSTM is a Recurrent Neural Network (RNN) architecture which is capable of doing sequential / time-series learning and predicting. It gets long-term dependencies by connecting previous information to the present task. We can obtain the parameters learned from dependent cascade model using LSTM and compare it to our proposed model.

Chapter 4

Bus Bridging in Post-event Crowd Diffusion: A Spatial Redistribution Approach

4.1 Overview

4.1.1 Motivation

In architectural and urban design community, there is a growing trend to design and build increasingly larger facilities that integrate diverse functions [55]. Examples of such facilities include stadiums, convention centers and airports. Operating such facilities with high volumes of human traffic is very challenging and needs to be carefully planned. Issues related to the operation of such facilities include, but not limited to, way-finding inside the facility, regular egress, and emergency egress. In particular, to serve the transportation needs of crowds moving into and out of such facilities, an important consideration is to integrate mass transit to the facilities.

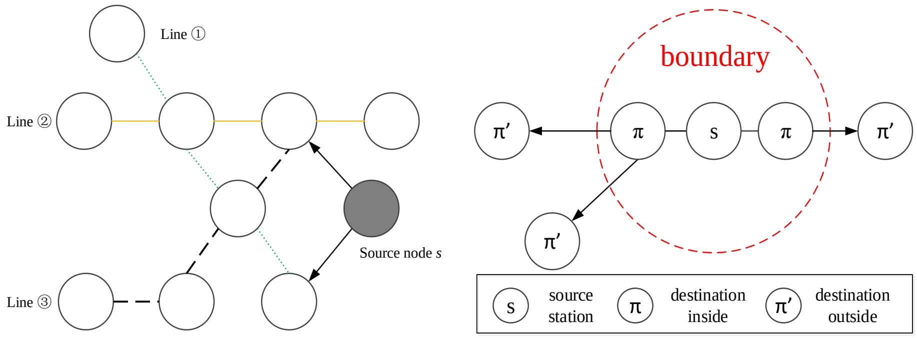
In this work the key motivation is to design a bus bridging service to complement mass transit during regular egress in order to minimize total travel time of crowds. While regular egress is predictable in both crowd volume and timing (the planner should know exactly how many people will be leaving the facility, and at what time), and all utilities can be assumed to be in perfect working condition (which contrasts the case of emergency egress, where the timing is uncertain, and some utilities could be faulty), the planning problem is still challenging. The major challenge in regular egress is to avoid bottlenecks and crowd buildups, which is hard to avoid since mass transit is designed to satisfy regular transport demands and not demand surges. A popular solution adopted by many planners is to complement mass transit with bus bridging services, yet despite the long history of using such services, optimizing its delivery has not received much attention; as a result, the design of bus bridging services for regular egress is usually ad hoc and static.

4.1.2 Problem Description

Our problem can be represented by a graph $\hat{G} = (\hat{N}, \hat{E})$ incorporating existing public transport service lines, where stations are denoted as nodes $\hat{N}$ and connectivities are denoted as directed links $\hat{E}$. An example can be seen in Figure 4.1(a), where each of the three lines is represented by a different line style. Stations along all lines are represented as hollow nodes in Figure 4.1(a), the node with ultra-high demand is shaded as node $s \in \hat{N}$. Note that node *s* is not necessarily connected to existing stations, and visitors at node s might need to find their ways to the closest station. This might be feasible for normal circumstances, yet when the demand is beyond planned capacity, this sudden inflow of demand might overwhelm the service provided at the nearby stations.

In this context, the bus bridging service between origin and destination is essentially a way to *add temporary routes* to the underlining network: if the existing routes between origin and destination do not have enough capacities to transport the allocated flow, a new route will be created; otherwise, no other routes will be generated in between. Noted that a route is composed of a set of connected arcs. Variable σ_τ^r denotes the deployment of bus service on route r and time τ. The total number of buses that can be deployed is bounded by the budget B.

If the current resource is enough, all of the commuters will travel via public



(a) An example of public transportation network with a node emitting ultra high demands.

(b) An example of destinations inside/outside the boundary, where the inside represent the bus service destination and outside represent their final destinations

Figure 4.1: Network structure and problem statement

transportation systems. Otherwise, a proportion of commuters has to find their own way getting home and hence we add a penalty when measuring their travel time. Commuters are characterized as a set of commodities based on their OD: there are divided into k groups and commuters of the same group share exactly the same origin and destination pair. We use variable $y^r_{k,\tau}$ to determine the number of passengers in group k that take bus service operate on route r at time τ and variable $\hat{y}_k$ to represent the unserved passengers in group k. As our major focus in this work is on the capacity planning, rather than scheduling, we assume passengers arrive in one shot, other than arriving in batches. Noted that the number of visitors that can be served is constrained by the bus services that have been deployed. Moreover, when we design the bus service in this model, we did not take congestions caused by bus service into account, which is one of the limitations in our model.

Therefore, our work is concerned with the following problem:
computing the optimal bus deployment plan $\boldsymbol{\sigma}$, that can distribute passengers over the map and in the meantime, miniminizes the total travel time for all passengers $\boldsymbol{y}$ with a fixed budget B.

4.2 The Crowd Management Model for Dynamic Bus Transit Service

To focus only on the part of graph where the bridging service can reach within *reasonable amount of time*, we define $N \subseteq \hat{N}$ to be the set containing only nodes that can be reached from node s within X minutes (X is empirically set to be large enough to contain all nodes we will ever consider, i.e.: $X = 30$mins). This region bounds the area that for us to do flow redistributions. Similarly, we define $E \subseteq \hat{E}$ to contain all edges between nodes in N. The reduced graph $G = (N, E)$ will be our focus for the rest of the section. In this work, we plan for commuters' trip on reduced graph G and define a station $\pi \in N$ to be the temporary destination inside the boundary, which is nearest to the final destination π'. In other words, we consider the problem of distributing flow as well as minimizing travel time from node $s \in N$ to temporary node π. Figure 4.1(b) illustrates an example of the mapping between stations inside and outside the boundary.

Formally, our problem is defined on a graph $G = (N, E)$, where N contains a set of nodes and E contains a set of arcs. Arcs are directed and each arc $(i, j) \in E$ is associated with the capacity $cap_{i,j}$ and arc cost $\delta_{i,j}$. Passengers are represented by multi-commodity flow as they are flow over the network with multiple pairs of origins and destinations. The set of commodities are given and each commodity $k \in K$ is defined by a tuple (s_k, π_k, d_k), where s_k is the origin, π_k is the destination and d_k is the quantity of commuters in the group. Each commodity k maintains a set of available route $R_k \in R$ and each route $r \in R_k$ is constituted of a set of arcs (i, j). We use $\alpha^r_{i,j}$ to represent if route r covers arc (i, j) ($\alpha^r_{i,j} = 1$) or not ($\alpha^r_{i,j} = 0$). The cost of a route r is the summation of cost for all arcs that belong to r, $\delta^r = \sum_{(i,j)\in r} \delta_{i,j}$.

Enumerating all routes for each group k is impossible as there are $\sum_{k=0}^{k=n} P^k_n$ ways going to the destination station from the source station, where n is the total number of stations in between. Besides, given the information of k commodities, it is in-obvious to tell efficient candidate routes for each k. Therefore, in this case, we apply the column generation procedure to identify good alternative bus routes for each commodity k with consideration of the demand d_k.

In the literature, Kaj *et al.* proposed the column generation procedure for multi-commodity flow in [21]. Column generation is a technique in operation research to solve large linear programming problem efficiently. They key idea for column generation is to add better solutions into the existing solution pool and improve the objective iteratively. In Kaj's work, he intends to solve the routing problems in a capacitated network that can send commodities efficiently. He contributes in proving column generation as a promising way for solving the extended model, which is NP-hard.

Although it was motivated to solve the problem related to telecommunication applications, this approach can be extended to other fields as well. We in this work extend the methodology in Kaj *et al.*'s work [21] and generate the candidate bus route with following assumptions:

- Each route $r \in R_k$ serves for commodity k and starts its service from origin s_k and ends at destination π_k.
- Each route r visits the selected arcs (i, j) in sequence and return to the origin directly after visiting the last node.

4.2.1 Candidate Route Selection

With above assumptions, we apply column generation procedure to generate bus routes. Restrictive master problem and pricing subproblem run iteratively to generate beneficial candidate routes until there is no improvement.

Restrictive Master Problem (RMP) We formulate the RMP problem in Table 4.1 which employs the following decision variables:

$$y^r \in R_k : \text{Quantity of flow in group } k \text{ that takes route } r \tag{4.1}$$

$$\bar{y}_k : \text{Quantity of flow in group } k \text{ not assigned to any route} \tag{4.2}$$

Table 4.1: RESTRICTIVE MASTER PROBLEM

$$\textbf{min} \sum_{k,r \in R_k} \delta^r \cdot y^r + \sum_k \bar{y}_k M$$

$$\textbf{s.t. } \bar{y}_k + \sum_{r \in R_k} y^r = d_k \ \ \forall k, r \in R_k \tag{4.3}$$

$$\sum_{r \in R_k} \alpha^r_{i,j} y^r \leq cap_{i,j} \ \ \forall i, j \tag{4.4}$$

$$0 \leq y^r \leq d_k \ \forall r \in R_k \tag{4.5}$$

$$0 \leq \bar{y}_k \leq d_k \ \forall k \tag{4.6}$$

The objective function is to assign all commodities to the network with minimal cost. M is a large number which assigns the penalty to the quantity commodity that is not able to be assigned. Constraint (4.3) guarantees the flow conservation where the total amount of flow in each commodity k is equivalent to the sum of assigned and unassigned flow. Constraint (4.4) limits the flow passing through each arc not exceed the arc capacity. Constraint (4.5) and (4.6) ensures the non-negativity and we make both decision variable $x^r, \bar{x}_k$ linearized.

Pricing Subproblem (PSP) ω, σ are defined to be the dual variables corresponding to constraints (4.4) and (4.3) in respective. In this case, all of the commodities would be assigned on the graph and thus optimality holds if the following restriction holds:

$$\sum_{i,j} \alpha^r_{i,j} \cdot \delta_{i,j} - \sum_{i,j} \alpha^r_{i,j} \cdot \omega_{i,j} \geq \sigma_k \tag{4.7}$$

Otherwise, a path r will be added into R_k as a column.

Therefore, we formulate the pricing subproblem in Table 4.2 with the objective function minimizing the reduced cost. We introduce the binary decision variable $z_{i,j}$ which indicates the arc (i, j) will be covered on a route if it takes the value 1 and 0 otherwise. Constraint (4.8) restricts the generation of duplicate route that exist in route set R_k. Constraint (4.9), (4.10) and (4.11) ensure that the amount of incoming and outgoing flow for all vertices must be the same except the start node s_k to end vertex π_k.

Table 4.2: PRICING SUBPROBLEM

$$\textbf{min} \sum_{i,j}(\delta_{i,j} - \omega_{i,j})z_{i,j} - \sigma_k$$

$$\textbf{s.t.} \sum_{(i,j)\in r} z_{i,j} \leq size(r) - 1 \;\; \forall r \in R_k \tag{4.8}$$

$$\sum_j z_{i,j} = \sum_j z_{j,i} + 1 \; \forall i = s_k \tag{4.9}$$

$$\sum_j z_{i,j} = \sum_j z_{j,i} - 1 \; \forall i = \pi_k \tag{4.10}$$

$$\sum_j z_{i,j} = \sum_j z_{j,i} \; \forall i! = \pi_k, s_k \tag{4.11}$$

$$z_{i,j} \in \{0, 1\} \tag{4.12}$$

Column Generation Procedure In column generation procedure, the RMP and PSP are solved iteratively until the reduced cost become positive and hence there is no further improvement can be made. Note that our RMP is for multi-commodity flow assignment problem and thereby the PSP generates alternative routes for each commodity k at one time. In this case, we denote set S to be the set of commodities that already got the good assignment and thus we only focus on generating routes for those groups that are not properly assigned.

The column generation procedure is shown in Algorithm 1, which starts with a basic route set R_k for each commodity k. Those basic routes are formed by the MRT lines with smallest travel time. By solving the restrictive master problem, the

Algorithm 1 Column Generation

Require: $cap_{i,j}, \delta_{i,j}, d_k$
function ColumnGen
  $R_k \leftarrow$ MRT routes
  for do
    $\omega, \sigma \leftarrow RMP(R)$
    for $k \in K$ **do**
      if $k \notin S$ **then**
        $r \leftarrow PSP(\omega, \sigma, R_k, k)$
        $\lambda_k \leftarrow \sum_{k,i,j}(\delta_{i,j} - \omega_{i,j})z_{i,j} - \sigma_k$
      if $\lambda_k \geq 0$ **then**
        $S \leftarrow S \cup k$
      else
        $R_k \leftarrow R_k \cup r$

dual variables ω, σ are obtained. With such dual variables, we solve the pricing subproblem for each commodity k. While the reduced cost of commodity k, λ_k is larger than 0, we keep generating new routes for this group. Otherwise, we stop generating new routes for it. We use set S to record the set of the commodity that is well assigned.

4.2.2 Bus Resource Allocation

Given the set of candidate route R_k for each commodity k generated in Section 4.2.1, the next phase is to plan for an optimal schedule with a limited number of buses and distribute passengers out of the congested region so that to minimize those commuters' travel time. As we generate routes for each commodity separately, there might be a case where route r for a commodity is part of the route r' for the other commodity. In this case, before planning for the buses, we first of all, organize those routes for all commodities by merging sub-routes to the complete routes. In the meantime, we generate variable $\gamma^r_{k,i,j}$ as the number of commuters in group k that is assigned to take route r which covers arc (i, j).

To schedule those available buses, we introduce the time horizon τ. Before a bus starts its service at time τ, commuters are queuing at station s. From the commuters' perspective, the total time taken during this process is comprised of the following components:

- Total travel time on route r;
- Waiting time for boarding the bus at s at time τ;

To plan for the bus deployment on each route with above considerations, we formulate the model in table 4.3 which employs the following decision variable:

$$\sigma_\tau^r : \text{Indicate employ a bus on route } r \text{ at time } \tau \text{ or not} \tag{4.13}$$

$$y_{k,\tau}^r : \text{Number of commuters in group } k \text{ take route } r \text{ at time } \tau \tag{4.14}$$

Table 4.3: RESOURCE ALLOCATION MODEL

$$\begin{aligned}
\textbf{min} \; & \sum_{k,r,\tau} (\delta^r + \tau) \cdot y_{k,\tau}^r + \sum_k \theta_k \cdot \bar{y_k} \\
\textbf{s.t.} \; & \sum_{r,\tau} y_{k,\tau}^r + \bar{y_k} = d_k \; \forall k && (4.15) \\
& \sum_{r,k} \gamma_{k,i,j}^r \cdot y_{k,h}^r \leq \sum_r \sigma_\tau^r \cdot \alpha_{i,j}^r \cdot Q \; \forall i, j, \tau && (4.16) \\
& 0 \leq y_{k,\tau}^r, \bar{y_k} \leq d_k && (4.17) \\
& \sigma_\tau^r \in \{0, 1\} && (4.18)
\end{aligned}$$

Objective function is constituted by two parts: (1). the total travel time on route r together with the commuters' waiting time when to board the bus departing at time τ; (2). the penalty for those commuters in each group k that are not able to rely on our planned bus to reach their destinations. Constraint (4.15) ensures the total demand in group k must be equivalent to the total demand that is served plus those are not served. Constraint (4.16) guarantees that, for each arc (i, j) and each time τ, the amount of commuters taken on arc (i, j) should not exceed the bus capacity that employed on this link. Q is the capacity associated with arc (i, j), either the capacity of an MRT arc or a bus arc.

4.3 Greedy Baseline

Besides the optimization based approach, in this section,we propose the greedy baseline to solve the crowd management problem in Algorithm 2.

Algorithm 2 Greedy Heuristic

Require: $\boldsymbol{d}, f_0, R_t, R_b, B$

function GREEDYBASELINE
    for $r \in R_t$ **do**
        $\Gamma \leftarrow \Gamma \cup (r, f_0)$
    $\delta = calTravelTime(Q, \Gamma)$
    $b = 0$
    while $b < B$ **do**
        for $\tau = 1..T, r \in R_b$ **do**
            if $(r, \tau) \notin \Gamma$ **then**
                $\Gamma' \leftarrow \Gamma \cup (r, \tau)$
                $\delta' = calTravelTime(Q, \Gamma')$
                if $\delta' < \delta$ **then**
                    $\delta \leftarrow \delta', \gamma \leftarrow (r, \tau)$
        if $\gamma ! = null$ **then**
            $\Gamma \leftarrow \Gamma \cup \gamma, b \leftarrow b + 1$
        else
            break

Input information includes demand of commuters $\boldsymbol{d}$, fixed frequencies f_0 of train service, train route set R_t, candidate bus route set R_b and limit budget B. The candidate bus route set R_b contains the direct link connecting start and end stations. In Algorithm 2, we denote Γ as the bus deployment plan, which is formed by tuples (r, τ) containing the information of route r and time τ. Before starting to search the strategy, Γ is initialized as incorporating the existing train services with known frequency f_0. Commuters are assigned under the initial deployment Γ and the total travel time τ is thereby measured. Variable b is used to record the number of available buses employed and it is initialed to be 0 at the beginning.

When b does not exceed the budget B, we iterate each time horizon τ and each bus route $r \in R_b$, forming a tuple $\gamma' = (r, \tau)$. A temporary deployment plans Γ' is set as the current deployment plan together with γ'. If Γ' can help disperse the large crowds and achieve the smaller travel time δ', we record the tuple and travel time and continue. This process ends with the tuple γ which decrease the travel time most. Consequently we include γ into the deployment set Γ and increase b by 1. Such procedure repeats until exceeding the budget. If adding additional bus deployment plan dose not help reduce the travel time, we break and jump out of the while loop.

4.4 Experiment

This section is organized as follows: firstly we compare the effectiveness of crowd distribution approach to the greedy baseline in section 4.4.1. After which, we do the sensitivity analysis over different population size in section 4.4.2. As our model can be generalized to deal with the large crowd due to train disruption, we compare the performance of our approach to a state-of-art disruption response model in section 4.4.3.

4.4.1 Comparing to Greedy Baseline

In this section, we show how our approach complements the mass transit services when the train is in perfect working condition. To concretely evaluate the performance, we compare the optimization-based approach described in Section 4.2 to the greedy baseline shown in Algorithm 2.

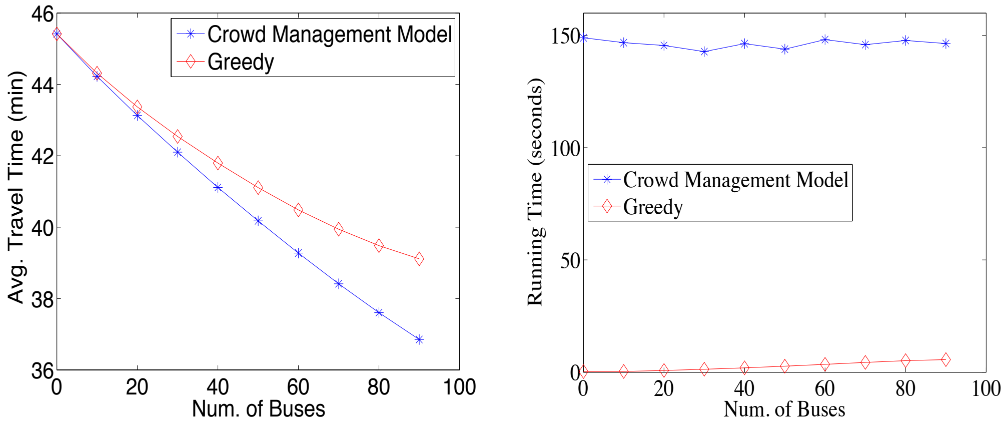
The experiment is conducted with the settings in Table 4.4. There are in total 25 stations shared by both MRT and bus networks. $50,000$ commuters are divided into 23 groups will be distribute from Singapore Sports Hub to their home locations. Regarding those commuters who are not able to board the bus or MRT, we make the penalty θ_k as the walking time from Singapore Sports Hub to destinations of the commuter group k.

4.4.1.1 Results

We provide performance and run time comparisons between the optimization approach and the greedy heuristic in Figure 4.2. Figure 4.2(a) plots the travel time

Table 4.4: Experiment setting

Item	Details
Num. of stations	25
Num.of BUS arcs	600
Groups of Commuters	23
Number of Commuters	50000



(a) Average travel time

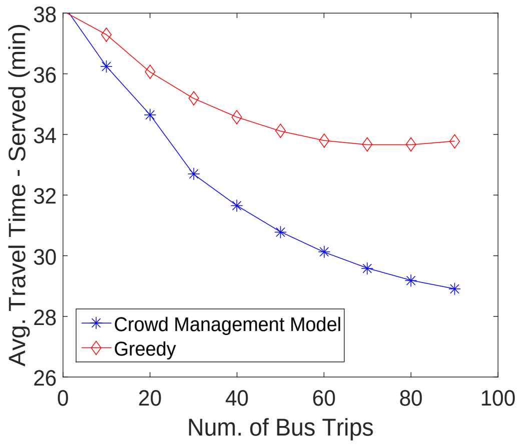
(b) Running time

Figure 4.2: Comparing to greedy heuristic

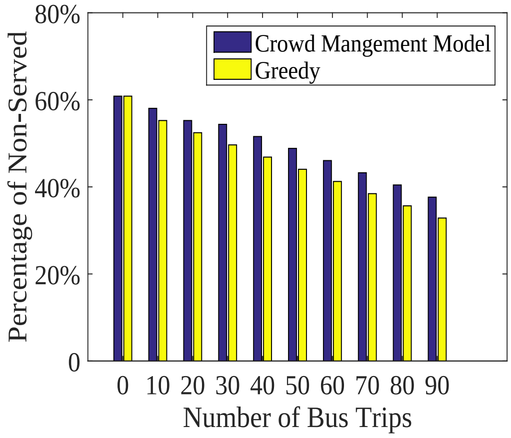
reduces due to the implementation of crowd management approach in comparison with the greedy baseline. Y-axis represent the commuters' average delay in travel time and x-axis denotes the number of available buses employed. We observe the average travel delay decreases when the number of employed bus increases for both methods. Intuitively, when fixing the number of buses deployed, smaller travel time indicates better performance. From the figure, our approach consistently performs better than greedy heuristic with 10 to 90 buses employed. Specifically, when the number of employed buses is 90, our approach can achieve 18.8% of the average travel time reduction. While for greedy heuristic, the reduction is 13.8%, which is 5% worse than optimization method.

We plot the running time for both methods in Figure 4.2(b). Apparently, greedy heuristic runs much faster than our approach. However, when looking at the running time for our approach, it is still considered affordable as it is around 150 seconds.

To have a comprehensive understanding of Figure 4.2 (a), we plot the average



(a) Travel time for served passengers



(b) Percentage of non-served passengers

Figure 4.3: Served and non-served

travel time for served passengers and the percentage of non-served passengers in Figure 4.3. Figure 4.3(a) indicates that our approach works better in serving the passengers who are able to be on board. This can be seen from the travel time reduction for our approach over greedy heuristic. The reduction achieved by Figure 4.3(a) is larger than it in Figure 4.2 (a). Figure 4.3 reports the percentage of non-served passengers for our approach and greedy heuristic. Together with the Figure 4.2(a), we observe that even the percentage of non-served passengers for our optimization based model is high, the travel time is lower. In other words, our model utilized better bus route to deploy its services and thus achieve lower travel time.

There are in total 25.2% of the commuters assigned to take the bus for both of the optimization-based approach and greedy baseline. To get an insight of the routes and commuter assignment, we plot the details in Figure 4.4 and 4.5. On both figures, lines with different line styles and color represent different routes and vice versa. Numbers associated with each link represent the proportion of commuters that adopt this one. When looking at the routes adopted by both methods, we observe that there are 4 common ones ($CC6-NE10$, $EW10-EW7$, $CC6-NE8$ and $CC6-NS24$) and we plot them with the same line style and color in both figures. The 4 routes are important in carrying commuters: for crowd management approach, the total percentage of commuters taking these routes is 10.1%; for greedy baseline, 18.2%

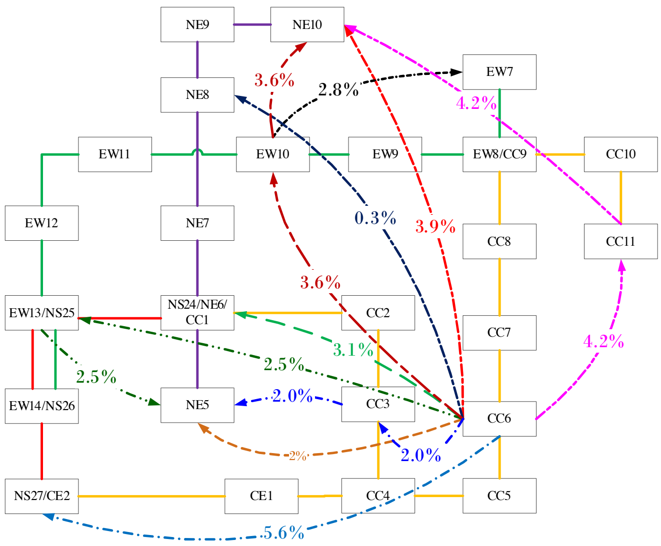


Figure 4.4: Routes of crowd management model

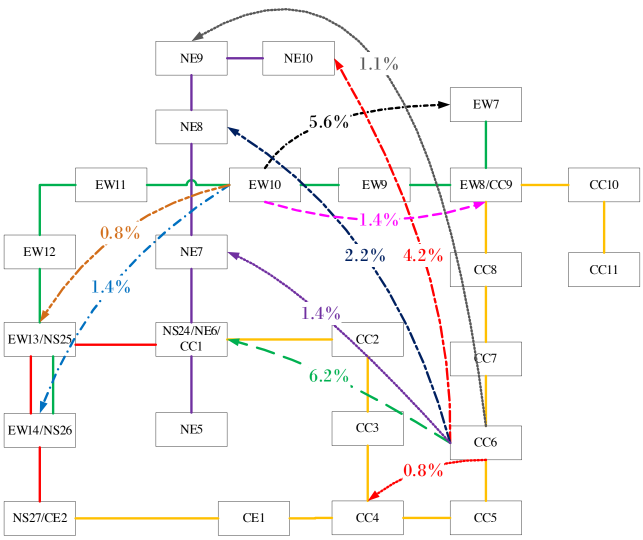


Figure 4.5: Routes of greedy heuristic

of the commuters is assigned on those routes.

Except the common routes, other routes adopted by optimization-based approach is different from greedy heuristic. In optimization-based approach, there are 9 routes starting from $CC6$ (a station near to the Singapore Sports Hub). While for greedy heuristic, despite concentrating on the nearest station $CC6$ to Singapore Sports Hub, it helps deal with the crowd around station $EW10$ as this station is within walking distance from Singapore Sports Hub. Essentially speaking, this result is due to the different ways of dealing with crowd of the two approaches:

optimization-based approach searches global optimal solution for all commuters in all groups and in the meantime, minimizing the total travel time. While the greedy baseline only focuses on commuters who are not able to board due to the limited resource and find local optimal routes for reducing the travel time for unserved commuters.

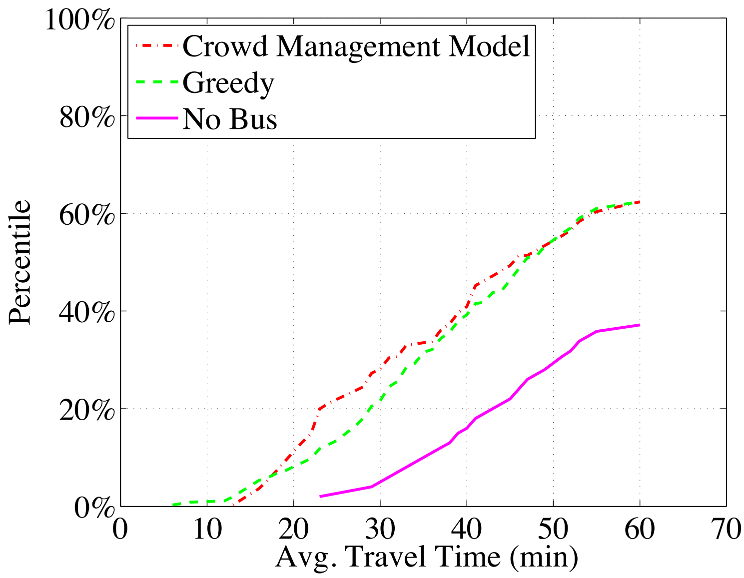


Figure 4.6: Percentile of arrivals

We plot the relationships between percentile of arrivals and average travel time in Figure 4.6. The solid line represent the baseline where there is no bus employed. The rest two lines with different line styles represent the results derived with different approaches. From this figure we observed that using the bus deployment derived by crowd management model, with average travel time is 30 minutes, 28.1% of commuters arrived at their destinations. However, with the same average travel time for greedy heuristic, 21.6% of commuters arrived. The baseline with no bus employed, only 6% of commuters arrived at destinations. From another aspect, 60% of commuters arrive at their destinations with average travel time around 53 minutes derived by both crowd management approach and greedy heuristic. In the end, almost 40% of the commuters reply on their own way to go home so they take their relative penalty with them. While for the baseline, where there is no external intervention on commuters' journey, only around 39% of the commuters arrived at home with average travel time 60 minutes.

4.4.2 Sensitivity to Different Population

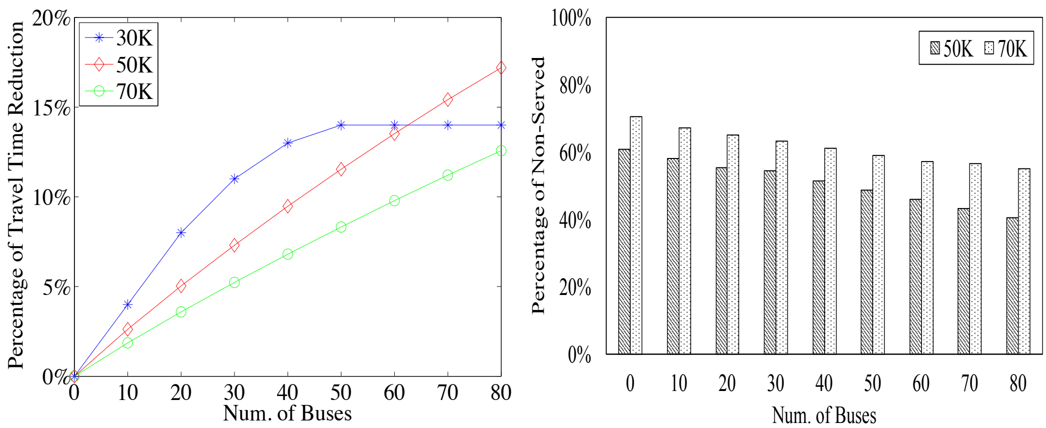
We now consider the travel time reduction due to the employment of buses with regarding to different population size. The travel time reduction Δ_i with i buses incorporated can be calculated as follows:

$$\Delta_i = \frac{\delta_0 - \delta_i}{\delta_0} * 100\% \tag{4.19}$$

where δ_0 is the original travel time with no bus incorporated and δ_i is the travel time with i bus incorporated. We plot relationship between Δ_i (y-axis) and i (x-axis) with different number of bus deployed for different population size in Figure 4.7(a). In this figure, three different line styles represent different population size and we observe that travel time reduction Δ_i get higher with increasing number of buses i for all of the three population size. When i reaches 80, population size $50,000$ gets the highest travel time reduction, 17.2%. While for $30,000$ and $70,000$ commuters, the reduction is 14.0% and 12.5%, respectively.

One interesting question is raised as the reduction Δ_i for $50,000$ is higher than that of $70,000$ when fixing i. Refer to Figure 4.7(b), this is because the demand surges of $50,000$ commuters is still considered affordable to the current public transportation system as almost 40% commuters can be served by public transportation system without any external intervention. With 80 buses incorporated, the percentage of the served commuters increases and is approaching 60%. While in terms of $70,000$ commuters, the current service system is significantly overwhelmed as over 70% commuters cannot be served without any buses. Even with 80 buses. this number slightly decreases to 55.2%, which implies that in the situation of $70,000$ commuters, a large portion of them are not able to be served and the therefore the penalty associated with each non$-$served commuter contributes to increase the overall average travel time.

Another interesting phenomenon occurs for population size $30,000$, where the reduction increases to 14% for 50 buses and keeps flat afterwards. This situation indicates that with enough buses, adding more resources does not help. MRT and buses come in sequence and such sequence make commuters wait at the station. Although taking MRT or bus helps reduce the total travel time for commuters, commuters may not afford the wait time if they have to wait for too long. In this case, commuters would rather prefer to find their own ways traveling home, which causes a relatively smaller travel time penalty: θ_k for commuters of group k. Consequently, the average travel time reduction keeps the same with enough buses incorporated.



(a) Travel Time Reduction with Different Population Size

(b) Percentage of Non Served Commuters

Figure 4.7: Sensitivity to population size

4.4.3 Comparing to Disruption Response Model

In the literature review section, we have briefly introduced the work proposed by Jin *et al.*[24] which presented a state-of-art optimization-based approach that responds to the disruption over massive transit train network by introducing the bus bridging services. This work is relevant to ours in the sense that it also adopts shuttle buses as a way to manage the large crowd that degenerates the current public transportation system. However, the major difference occurs as the model proposed by Jin *et al.* must be in the context of train disruption and it does not work out with fully connected train network. This is due to the design of column generation when deriving the candidate routes: with perfectly connected train network, existing routes are efficient enough and there is no potential for objective value in restricted master problem to improve. Thus, dual prices for each of the relevant constraints in the master problem is 0 which makes the objective value of the subproblem 0. In this case, no effective bus routes can be added into the routes set.

While our model is more general as we are able to deal with the large crowds regardless of the connectivity of existing network. Unlike Jin *et al.*'s work [24] , which designs column generation procedure by taking network connectivity into consideration, we make routes capacity as a key indicator to decide whether adding more routes or not: while the current network does not have enough capacity to transport all commuters, new routes are adding in; otherwise, the column generation procedure stops.

To show the performance of our model under the train breakdown situation, we next compare the results for the two approaches in context of train disruption.

4.4.3.1 Disruption Case

The specific parameters of the disruption scenario are shown in Table 4.5. The network contains 31 MRT nodes. Out of which, 8 links were broken down and hence the capacities relevant to these links are set to 0. Each MRT station has a nearby bus station but as there are some interchange MRT stations next to each other so that we only set one bus station among that area. In the end, there are 25 bus stations settled and we assume that buses can travel between any pairs of stations. Commuter information is derived from EZlink data set and 26 groups of commuters are affected by such disruption.

To compare the two works, we set the common parameters as follows:

- Standard bus bridging services are running between pairs of disrupted stations with the same frequency as the MRT.
- Bus capacity: 140

Disruption	
Num. of affected links	8
Disrupted links	CC3 – CC4 NE6 – NE7 EW11– EW12 EW12 – EW13
Network	
Num. of MRT stations	31
Num. of BUS stations	25
Num.of BUS arcs	600
Groups of Commuters	26

Table 4.5: Disruption parameters

- Bus frequencies are between 1 $\sim$ 6 mins.
- Penalty for commuters who are not able to board a bus: 50 min.
- The number of buses starting simultaneously from each station should not exceeding 3.

In real world, there are two major types of disruption: unexpected and planned.

. **Unexpected disruption** happens in a sudden. For instance, in some cities like Singapore, train service always breaks down unexpectedly and commuters have no idea of such information until they arrive at one station just before the disruption. In this situation, commuters are not able to plan their journey at the time they depart. Instead, they have to find alternative ways that can bridge the broken link only when they reach the affected station. In Figure 4.8 for instance, realizing that links between e' and e are still connected, commuters in this scenario are probably trying to find the alternative bridging services between s' and e' and after that reply on the rest non-affected metro system from station e' onwards to their destinations.

. **Planned disruption** refers to the case where metro system is under maintenance and notifications of affected links and timing are revealed to commuters beforehand. For example, in some cities such as London and Boston, the metro systems periodically close due to regular maintenance. Commuters are aware of the current transportation network structure and are expected to find alternative ways from origin to destination before departure in order not to be affected by such disruption. For example, in Figure 4.8, a group of commuters have their original starting station s and destination e, which covers a disrupted link (s', e'). Noticing that there will be a link closure between s' and e', commuters in this group are going to plan their alternative ways from s to e before they depart. In this case, commuters travel demand is from s to e.

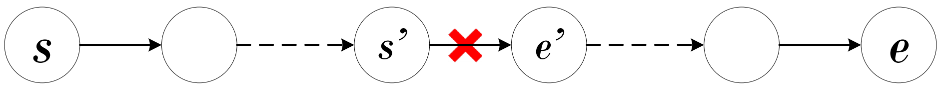


Figure 4.8: Disruption information

Although in above two situations, commuters react to the different types of disruption in different ways, the commuters demand scenario are exactly the same. In the following sections, we experimentally show the performance of these two models to response to the above situations. For simplicity, we call the first scenario *unexpected disruption* and similarly, the second one *planned disruption*.

4.4.3.2 Unexpected Disruption

In this section, we keep the same travel demand but alter the commuters' start station and end station for their trips' planning as in this context they might plan their trip in a different way. Refer to Figure 4.8, start station s' is defined as the first station just before the disruption and end π' the first station just after the disruption.

We compare the relationship between commuter travel delay and the number of buses deployed for both works and report the results in Figure 4.9. As can be seen, increasing the number of deployed buses reduces the average delay for all commuters for both methods. A key observation is that when number of deployed buses are larger than 20, travel delay of Jin *et al.*'s approach goes flat while our approach keeps going down. This is because commuters in Jin *et al.*'s approach are mostly assigned to take the standard routes as they do not consider the route capacity when do the assignment. However, the number of buses scheduled on the standard routes are very limited. Thus adding more buses does not help. We further clarify this in the following paragraphs.

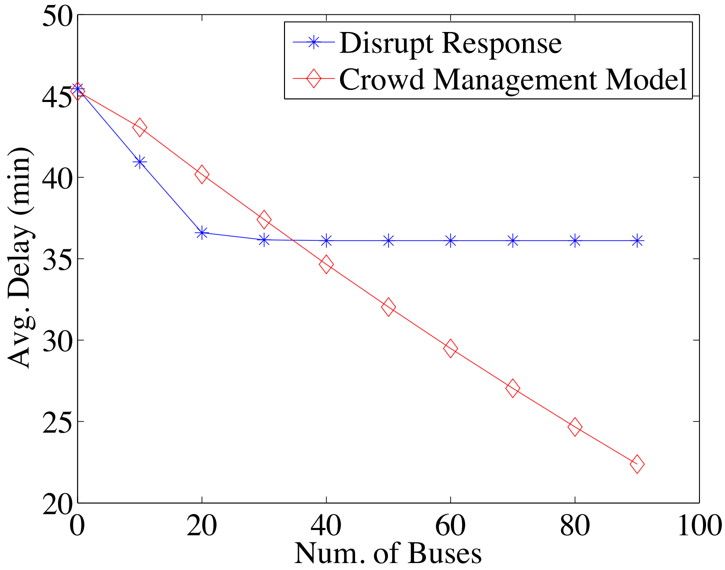


Figure 4.9: Average delay

Different column generation process in both works lead to the different commuters assignment on routes. In Jin *et al.*'s work, the column generation procedure does not consider the link capacity limit when assigning commuters. Hence, most commuters are assigned only to the routes with the least travel time – the standard bus bridging routes that connect two neighboring disrupted stations. While in our approach, noticing that existing routes capacity cannot fulfill commuters' travel demand, our column generation process creates more diverse routes to disperse the commuters flow to fulfill all commuters' travel demand. Figure 4.10 and 4.11 plots the routes and assignment of commuters derived by column generation procedure for both Jin *et al.*'s method and our approach in respective. For simplicity, we only plot those routes that has commuters assigned on it. Those routes directly bridging the disconnected links are standard bus routes (refer to the second row in legend). Routes of other styles are additional routes added by column generation procedure.

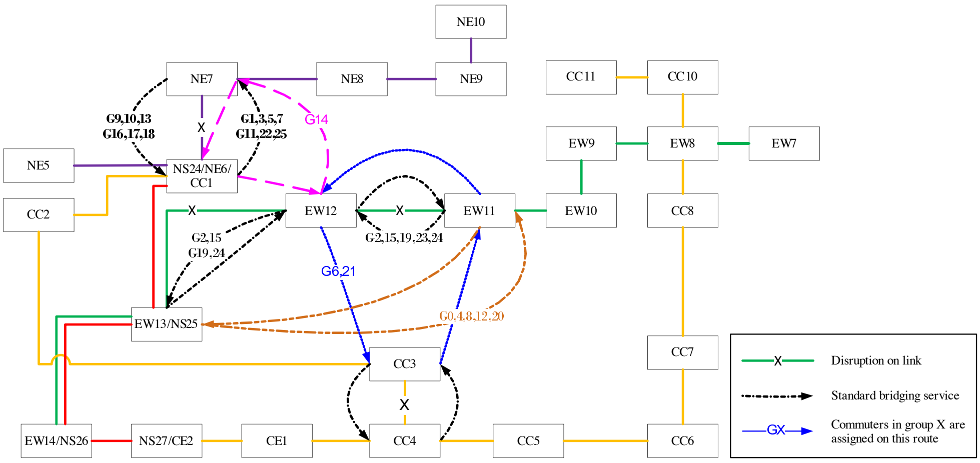


Figure 4.10: Routes and assignment of Jin *et al.*'s approach

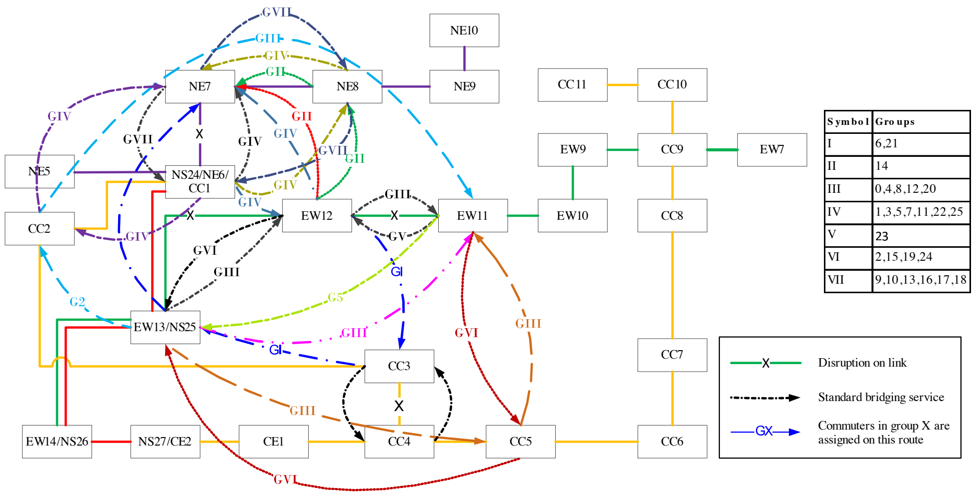


Figure 4.11: Routes and assignment of crowd management model

Comments on the routes indicate the group of commuters are assigned to take this route, e.g.: G4,G5 on the link connecting station EW12 and EW11 indicates that this link will be taken by commuters in group 4 and 5. As can be seen, in Figure 4.10, there are 3 additional routes added besides the standard bridging routes. 38.8% of the commuters are assigned to take these routes and the rest 61.2% rely on standard bridging services. While in Figure 4.11, only 34.9% of the commuters takes standard services to travel and the rest 65.1% of commuters take the additional 16 routes added by column generation.

The number of buses scheduled on the standard routes are very limited. We explain this point with the Figure 4.12 and 4.13, which plot the number of buses

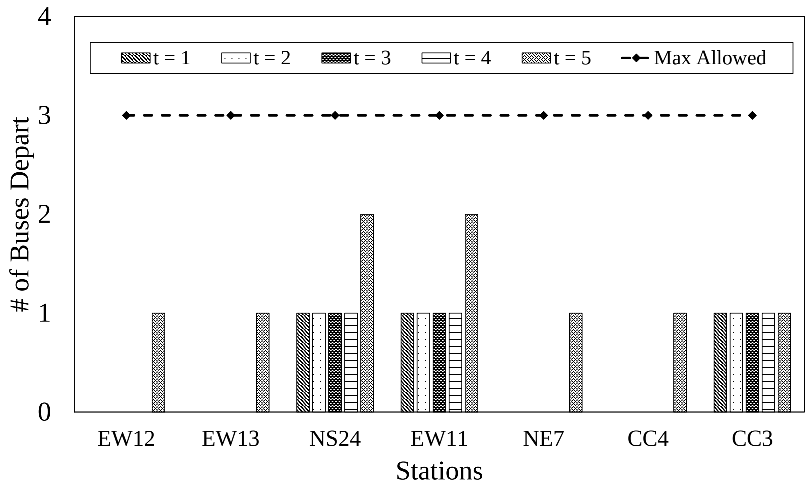


Figure 4.12: Routes and assignment of Jin *et al.*'s approach

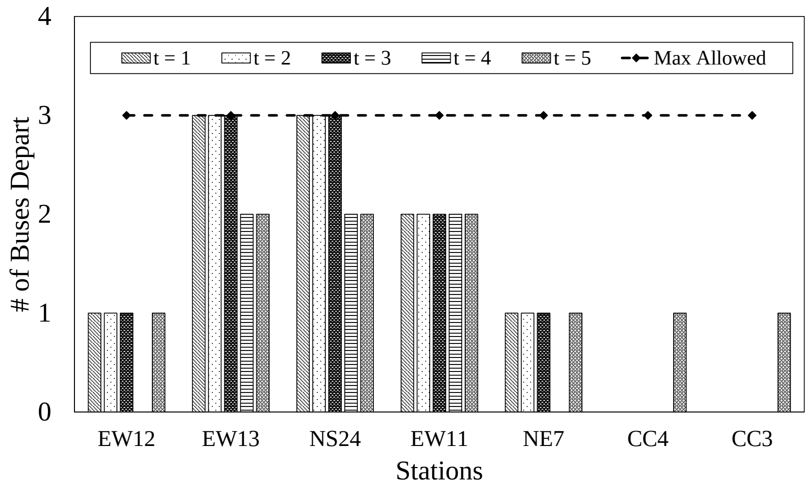
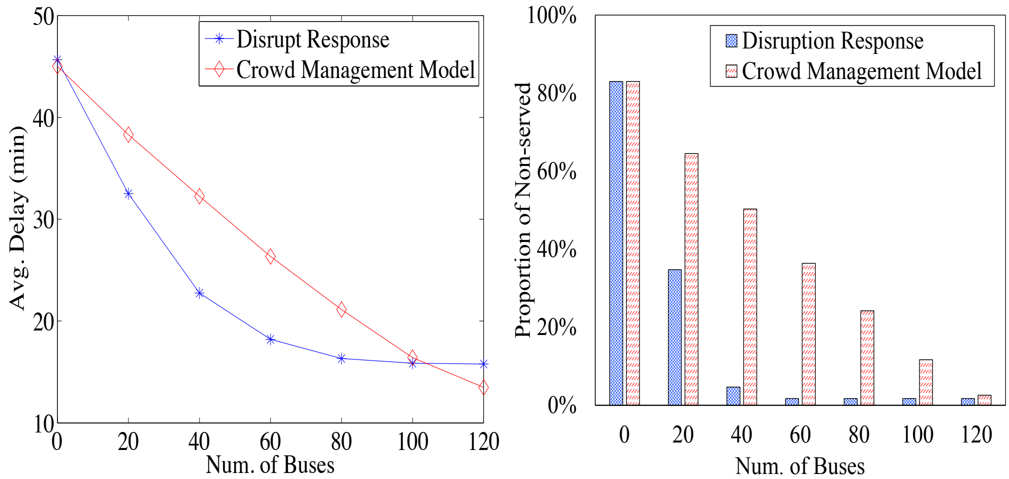


Figure 4.13: Utilization of stations of crowd management model

depart from various stations for both approaches in X-axis denotes a list of stations and y-axis plot the maximum number of buses that can start simultaneously. For simplicity, we only plot the status of those bus stops in the first 5 time unit as standard bridging services only provide services in intervals in $5, 10, 15 ...$. To make it a fair comparison, we make the number of buses departing from the same location not exceed 3. According to Jin *et al.*'s approach, only 3 additional routes are utilized besides the standard response. Those routes starts from $NS24, EW11$ and $CC3$ in respective. As there is a common constraint for both model: only one bus can be deployed to a route at one time. Therefore, the number of buses starting from each of the 3 stations is 1 before $\tau = 5$. At $\tau = 5$, there will be additional buses serving for standard routes. In this model, most commuters are assigned to the standard bus routes, where the frequencies are fixed, adding too many buses on other routes does not increase the volume of commuters to be transported. Therefore, the travel delay does not continue to drop with more than 40 buses. Refer to Figure 4.12, the departing buses of each bus stop is far below 3, which implies that the utilization rate of resources is actually small. While for our approach, commuters are dispersed to a variety of routes and adding more buses to various routes help to transport commuters. According to Figure 4.13, we make many buses departing from the same station serving different routes. For example, in the first 3 time units, the number of buses starting from $EW13$ and $NS24$ reaches the top limit of 3, which indicates that we make a high utilization rate of the bus stop resources. As a consequence, travel delay keeps decreasing with more than 40 buses.

4.4.3.3 Planned Disruption

In case of planned disruption, we plot the relationships between commuter travel delay and the number of buses deployed in Figure 4.14. As can be seen, increasing the number of deployed buses reduces the average delay for all commuters for both methods. A key observation is that with the smaller number of buses employed, disruption response get smaller travel delay than our approach. However, our approach get better performance with enough buses employed. Such situation can be explained with the help of Figure 4.14(b). In this figure, we observe that the proportion of commuters that cannot board on bus derived from our model is significantly larger than that proposed by disruption model with number of buses ranged $20 \sim 80$. However, with enough buses (≥ 100) available to deploy, the proportion of non-served commuters for both methods get closer. In such case, the average delay in Figure 4.14(a) for both model become closer as well. With more than 120 buses, our approach even get better performance. Regarding to the proportion of non-served commuters shown in Figure 4.14(b), an interesting question is raised: why the non-served commuters derived by the disrupted response approach is significant smaller than the one derived by our work with the same number of buses



(a) Average Delay (b) Proportion of Non-Served

Figure 4.14: Comparing different number of buses

employed. Such results can be explained by the different ways of doing column generations. For both works, column generation procedure searches for beneficial candidate bus bridging routes and in the meantime, assign commuters to beneficial routes. Such assignment is going to affect commuters' total travel delay. For instance, there will be no travel delay on MRT links whereas bus routes creates the travel delay. Furthermore, MRT links are able to transport many commuters while bus links have capacity limitation.

In Jin *et al.*'s work [24], commuters only take certain legs of the bus routes. The rest of their trips rely on non-disrupted MRT links. Therefore, closure of a MRT link leads assignment of commuters to various alternative bus legs that can bridge such link. In this way, the total travel delay only consists of the extra travel time on those bridging bus legs. However, in terms of our approach, if commuters are assigned to a route, they must take all links along it. In another word, for routes containing both MRT and bus links, the capacity of such route is limited by the bus lines, which has smaller capacity. In fact, the disrupted MRT networks are not able to fulfill commuters demand and we have to bridge them by bus legs, and the capacity bottleneck limits the efficiency of the transportation network. In this way, the non-served commuters derived by our model is larger than Jin *et al.*'s model. However, from another perspective, our model is not designed to handle the disruption case. We mainly focus on complementing the MRT services in normal situation without any MRT breakdown and we would not have this capacity bottleneck then.

4.5 Conclusion

Operating large facility in the urban city is challenging in many aspects. Distributing the passengers using public transportation system out of the large facilities is one of the hot spots. In this work, we target on minimizing the negative effect of the large crowds on the existing infrastructure. In doing so, we propose a two-phase optimization based approach to find efficient routes for commuters and schedule buses over the generated routes. This approach is general in the sense that it works out in both non-disrupted and disrupted transportation networks. In the normal situation where the network is in perfect working condition, we propose a greedy baseline and experimentally show the comparison to the optimization approach. Results show that the optimization approach is able to save 5% of the average travel time compared to the greedy baseline. We demonstrate the usage of our approach by comparing it to a state-of-art work handling the disruptions. We conduct the experiment with the configuration of two types of disruption, expected and unexpected. Our optimization model can obtain better performance when the disruption is totally unexpected.

Chapter 5

Integrated Bus and Bike Sharing Services for Last-mile Commute: A Spatial Redistribution Approach

5.1 Overview

5.1.1 Motivation

Singapore is a world-class city with many foreign companies keen on establishing their business in the region. The government built many business parks in many different locations over the past decade. In some of the above business parks, there is only limited number of MRT stations (in most cases, one station) taking passengers from other areas of the city to the business park. During morning and evening peak hours, demand surge occurs at those major MRT stations. Moreover, there is limited services provide the service from MRT stations to passengers' final destinations. In this case, operators need to carefully design temporal transportation services that helps disperse the crowds during peak hours and send passengers to their office destinations within a fixed budget.

The key problem we studied in this work is to determine the best portfolio assignment that can plan over a mixture of transportation modes, in this work, we take bus and bike sharing services as an example. The key challenge in this work is to connect the two independent problems together and leverage the trade-offs between budget assignment and passengers assignment.

5.1.2 Problem Description

Our problem is defined on an underlying network $G = (N, E)$, shown in Figure 5.1. Node set N contains important nodes including transportation hubs, road crossings and office buildings. Directed arc set E denote the driveways which are connected by crossings. Buses follow driveways to travel and each bus route r is formed by a

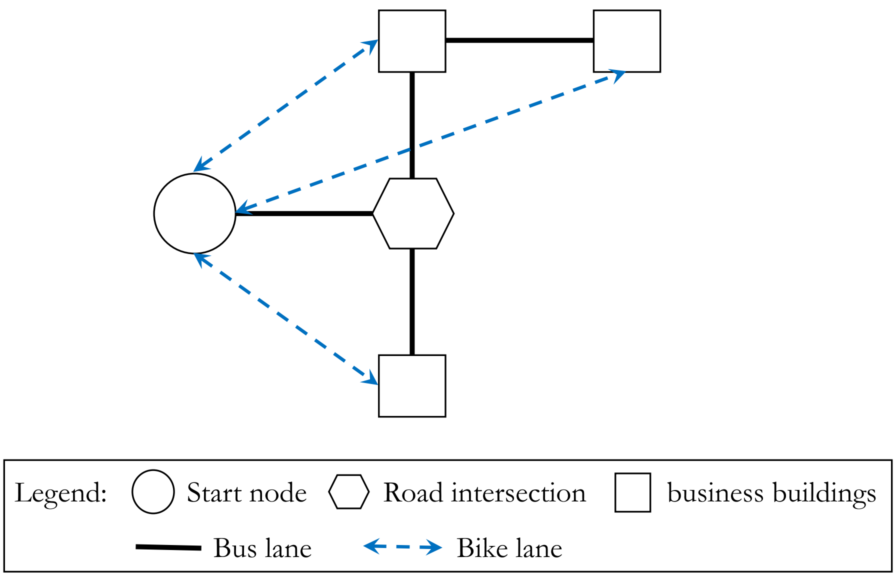


Figure 5.1: Network structure

sequence of arcs. A route must be simple, in other words, the repetition of nodes is not allowed. A route $r \in R_k$ serving passengers in group k, is defined as a path making start node s to destination π_k reachable. A route r covering several OD pairs can be shared by multiple groups of passengers. At different time of a day, bus travel time δ^r on route r is different and this information can be extracted from Google map. From Figure 5.1 we observe that routes of bikes are directly linked between stations. In fact, bike lanes can be flexible, and both driveways and sidewalks may be utilized according to passengers' preferences. The riding time between a pair of stations can be approximated as the average travel time for several routes that connects them.

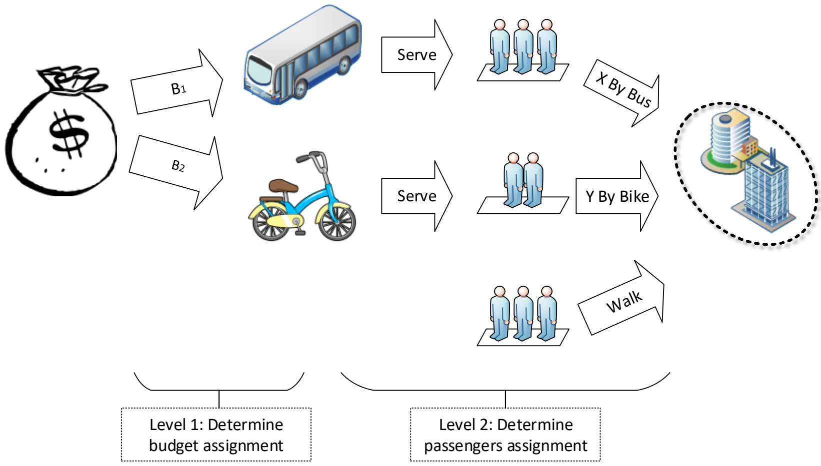


Figure 5.2: Problem structure

We consider providing a mixture of transportation options including both bus and bike services at a business park to distribute the demands and support the last-mile commute. The primary goal of the operators is to optimize their portfolio when implementing the services. Specifically, the major concern is to optimally split a fixed budget B to different transportation modes so that can achieve the desired purpose. Regarding passengers, their travel convenience can be quantified using the travel time. Shorter travel time offers a more convenient service to passengers. Therefore, this problem is divided into two steps indicated by Figure 5.2. In the first step, operators specify the budget B_1 to maintain bus services and $B_2 = B - B_1$ to operate bike sharing service. With the assigned budget to build infrastructures, X passengers are allowed to take the buses, Y passengers can take the bike, and the rest passengers have to walk to travel shown on the second step.

Modeling passengers' choice: Number of passengers served by each mode is highly related to how this mode is designed. We adopt the choice function $\beta_{i,j}$ introduced by [19] to describe passengers' preferences which has the following form:

$$\beta_{i,j} = \begin{cases} \exp^{-\frac{\hat{d}_{i,j}}{\hat{r}}} & \hat{d}_{i,j} \leq \hat{r} \\ 0 & \text{otherwise} \end{cases} \tag{5.1}$$

Equivalently, $\beta_{i,j}$ denotes the probability of passengers going from location i to j depends on the distance $\hat{d}_{i,j}$ between them. The higher the $\hat{d}_{i,j}$, the lower the $\beta_{i,j}$. When the distance between i and j is over a threshold $\hat{r}$, $\beta_{i,j} = 0$ and implies that passengers are not intend to go.

We incorporate choice function $\beta_{i,\pi}$ for both bus and bike services. In terms of bus mode, we assume whether passengers alighting at a station i depends on the distance of i to passengers' destination π. Smaller distance provides higher probabilities for passengers to get-off. In terms of bike mode, we assume whether passengers prefer to take bike mode depends on the distance from their origin and destination. If the distance is too far, passengers would choose to take other transportation modes rather than taking bikes.

The decision making problem in this work is to find the best trade-offs between investing different budget between bus and bike services so that can save passengers' travel time. Therefore, our work is to solve following problem:
Computing the optimal budget assignment B_1, B_2, which minimize the travel time of all the visitors, including those choose to take buses X, bikes Y and non-served passengers.

5.2 Designing Bus Services

To design the bus service inside the business districts, we have two major issues to consider: (1). identify good candidate routes; (2). determine the best bus deployment over different routes. We present the two parts in Section 5.2.1 and 5.2.2, respectively.

5.2.1 Generating Bus Routes

A bus route defined in this work is a path starting and ending at the source node, composed by a list of arcs on the graph. Enumerate all possible paths on the graph is costly and therefore, we employ column generation [57] technique to identify good bus routes. In which process, Restricted Master Problem (RMP) and Pricing Sub-Problem (PSP) run iteratively to derive good paths until there is no improvement.

5.2.1.1 Restricted Master Problem

In this section, we seek to find a set of efficient route R to serve passengers. Efficiency can be considered in many ways, such as the route with smaller travel time and so on. In this work, each route has a related travel time δ^r and consider the route with smaller δ^r more efficient. We introduce decision variable $z^r \in \{0, 1\}$ to represent whether we adopt route r or not. Hence, the objective function is:

$$\textbf{minimize} \sum_{r \in R} \delta^r \cdot z^r \tag{5.2}$$

We illustrate the problem with following constraints:

1. **Demand coverage**: Each route has a sequence of intermediate stops and if a station is within the walking distance to passengers' destination, we consider this route cover passengers' demand. We must ensure the demand coverage holds for each group of passengers:

   $$\sum_{r \in R} w_k^r \cdot z^r \geq 1 \ \ \forall k \in K \tag{5.3}$$

   We denote N_k as a list of stations which can serve passengers demand in group k, i.e. node $i \in N_k$ if $\beta_{\pi_k, i} \geq \phi$. In other words, if the probability of passengers walking towards their final destination π_k from node i is equal to or greater than a threshold ϕ, i belongs to N_k. We have $w_k^r = 1$ if route r contains nodes in N_k. Constraint (5.3) therefore indicates that at least one route r designed covers the demand in group k.

2. **Bound the utilization times**: We use the following constraints to denote that

the number of routes that are utilized must be within a tolerate range.

$$\sum_{r \in R} z^r \leq U \tag{5.4}$$
$$\sum_{r \in R} z^r \geq L \tag{5.5}$$

5.2.1.2 Pricing Subproblem

From the Restrictive Master Problem, we obtain dual variables associated with constraint (5.3), (5.4) and (5.5) as: λ_k, μ_i and ω. Hence, the reduced cost for each route r is:

$$c^r = \delta^r + \sum_{i \in N} g_i^r \cdot \mu_i - \sum_{k \in K} w_k^r \cdot \lambda_k - \omega \tag{5.6}$$

If reduced cost is negative, there is potential improvement for route r in Restrictive Master Problem. Otherwise, the Restrictive Master Problem is optimal. The following objective function (5.7) minimizes the reduced cost.

$$\textbf{minimize} \quad \sum_{i,j \in N} (\delta_{i,j} + p_j) z_{i,j} + \sum_{i \in N} \mu_i \cdot g_i - \sum_{k \in K} \lambda_k \cdot w_k - \omega \tag{5.7}$$

The node and demand coverage issues are captured by decision variables g_i and w_k in Equation (5.7). $g_i \in \{0, 1\}$ is to indicate whether node i is incorporated in a route and $w_k \in \{0, 1\}$ denote if the demand in group k can be served by this route.

Pricing subproblem generates a route starting and ending at source station s at a time. We use auxiliary variable $z_{i,j} \in \{0, 1\}$ as an indicator of whether arc (i, j) is included. The following constraints help form a valid route.

1. **Flow conservation**: In this context, each node can be visited at most once

and the inflow and outflow must be exactly equal.

$$\sum_{j \in N} z_{i,j} \leq 1 \ \ \forall i \in N \tag{5.8}$$

$$\sum_{i \in N} z_{i,j} \leq 1 \ \ \forall j \in N \tag{5.9}$$

$$\sum_{j \in N} z_{i,j} = \sum_{j \in N} z_{j,i} \ \ \forall i \in N \tag{5.10}$$

2. **Include source node**: A valid route must start and end at source node.

$$\sum_{j \in N} z_{s,j} = 1 \tag{5.11}$$

Constraint (5.11) ensures exactly 1 arc starting from source node s. Together with Constraint (5.10), there is exactly 1 arc ending at source node s as well, making this route a circle.

3. **Demand coverage**: We must make sure the demand coverage decision variable w_k is consistent with the routes formed using $z_{i,j}$.

$$w_k \leq \sum_{i \in N, j \in N_k} z_{i,j} \ \ \forall k \in K \tag{5.12}$$

$$w_k \cdot M \geq \sum_{i \in N, j \in N_k} x_{i,j} \ \ \forall k \in K \tag{5.13}$$

N_k represent the set of nodes that can serve demand group k, Specifically, if the probability of walking from a station $i \in N_k$ to the destination of k is equal to or greater than a threshold, i.e. $\beta_{i,\pi_k} \geq \phi$, we include i in N_k. $\sum_{i,j \in N_k} z_{i,j}$ denotes the inflow into demand node $j \in N_k$. If there is at least 1 arc connecting the demand node $j \in N_k$, the demand in group k is satisfied, $w_k = 1$. Otherwise, $w_k = 0$.

4. **Node coverage**: Once a node has outgoing edges included, it must be covered

in this generated route.

$$g_i = \sum_{j \in N} z_{i,j} \quad \forall i \in N \tag{5.14}$$

5. **Bounding the number of arcs included**: We use the following constraints to limit the number of arcs in a route not exceeding L_1 and the travel time of a route not exceeding L_2.

$$\sum_{i \in N, j \in N} z_{i,j} \le L_1 \tag{5.15}$$

$$\sum_{i \in N, j \in N} \delta_{i,j} \cdot z_{i,j} \le L_2 \tag{5.16}$$

When we generate a route of a circle, sub-tour may exist. To eliminate the sub-tour, we add Constraint (5.17) introduced by Jin *et al.* [23].

$$\sum_{i,j \in S^t} (1 - z_{i,j}) \le 1 \tag{5.17}$$

where S_t is the set of arcs belongs to the sub-tours in the t^{th} iteration.

The Restrictive Master Problem and Sub-Pricing Problem are illustrated in Table 5.1 and Table 5.2.

Table 5.1: Restrictive Master Problem

$$\textbf{min} \quad \sum_{r \in R} \delta^r \cdot z^r$$

$$\sum_{r \in R} w_k^r \cdot z^r \ge 1 \quad \forall k \in K \tag{5.18}$$

$$\sum_{r \in R} z^r \le U \tag{5.19}$$

$$\sum_{r \in R} z^r \ge L \tag{5.20}$$

Table 5.2: Subpricing Problem

$$\mathbf{min} \quad \sum_{i,j \in N} (\delta_{i,j} + p_j) z_{i,j} + \sum_{i \in N} \mu_i \cdot g_i - \sum_{k \in K} \lambda_k \cdot w_k - \omega$$

$$\sum_{j \in N} z_{i,j} \leq 1 \;\; \forall i \in N \tag{5.21}$$

$$\sum_{i \in N} z_{i,j} \leq 1 \;\; \forall j \in N \tag{5.22}$$

$$\sum_{j \in N} z_{i,j} = \sum_{j \in N} z_{j,i} \;\; \forall i \in N \tag{5.23}$$

$$\sum_{j \in N} z_{s,j} = 1 \tag{5.24}$$

$$w_k \leq \sum_{i \in N, j \in N_k} z_{i,j} \;\; \forall k \in K \tag{5.25}$$

$$w_k \cdot M \geq \sum_{i \in N, j \in N_k} z_{i,j} \;\; \forall k \in K \tag{5.26}$$

$$g_i = \sum_{j \in N} z_{i,j} \;\; \forall i \in N \tag{5.27}$$

$$\sum_{i \in N, j \in N} z_{i,j} \leq L_1 \tag{5.28}$$

$$\sum_{i \in N, j \in N} \delta_{i,j} \cdot z_{i,j} \leq L_2 \tag{5.29}$$

5.2.1.3 Column Generation Procedure

The column generation procedure adopted in this work is presented in Algorithm 3.

Algorithm 3 Column Generation Procedure

```
procedure ColumnGen
    R ← R ∪ R_0
    while do
        μ, λ, ω ← RestrictiveMasterProblem(R)
        r ← PricingSubproblem(μ, λ, ω)
        while sub-tour exist in r do
            add Constraint (5.17)
            re-solve PricingSubproblem(μ, λ, ω)
        R ← R ∪ r
        if rc ≥ 0 then
            break
```

We first initialize the route set R by incorporating a set of feasible route R_0. After which, the Restrictive Master Problem and Pricing Subproblem run iteratively until reduced cost $rc >= 0$, indicating no better routes can be generated.

5.2.2 Determining Bus Deployment

Given the set of bus routes generated by column generation approach, the next step is to determine the optimal deployment of buses over each route.

To distribute the passengers over the spatial horizon, the key concern for operators is to determine the number of buses to rent for each route and the number of service rounds to be provided over each route. Therefore, we introduce decision variable $h^r \in \mathbb{N}$ to denote the number of vehicles to rent or buy for route r and $f^r \in \mathbb{N}$ to represent the deployed round of services for route r. A good deployment plan is able to give more benefit to passengers under a fixed budget.

Passengers' benefit can be interpreted in many ways, such as the shortest travel time for all passengers, the expected number of passengers can be served and so on. In this problem, we quantify the benefit as the shortest travel time for all passengers. We use variable $\hat{y}^r_{k,i} \in \{0, 1\}$ to represent whether passengers in group k choose to take r alight at station i and $y^r_{k,i} \in \mathbb{N}$ to denote the number of passengers in group k choose to take r alight at station i. For those passengers in group k who are not able to be served, we use variable $\bar{y}_k \in \mathbb{N}$ to represent.

We therefore propose the objective function to minimize the average travel time for all passengers as follows:

$$\textbf{minimize} \quad \sum_{r\in R, i\in N_r, k\in K} (\delta^r_i + \bar{\delta}_{i,\pi_k}) \cdot y^r_{k,i} + \sum_{k\in K} \bar{\delta}_{s,\pi_k} \cdot \bar{y}_k \tag{5.30}$$

δ^r_i is the on bus time when passengers choose to take bus serving route r and alight at station i and $\bar{\delta}_{i,\pi_k}$ represent the walking time from station i to destination π_k. Thus the first term represent the travel time for those passengers that are served and the second term is the travel time for those unserved passengers.

We address the concerns of this problem using following constraints:

1. **Set upper bound for served passengers according to their choices**: A route

contains a list of intermediate stops and the distance between passengers' destinations and bus stops affects passengers' alighting choices. Shorter the distance leads to higher incentives for passengers to alight, which sets the upper bound on the amount of passengers taking a route. We represent it in following constraint:

$$y^r_{k,i} \leq \beta_{\pi_k,i} \cdot d_k \cdot g^r_i \cdot \hat{y}^r_{k,i} \quad \forall r \in R, i \in N_r, k \in K \tag{5.31}$$

$$\sum_{i \in N_r} \hat{y}^r_{k,i} \leq 1 \;\; \forall r \in R, k \in K \tag{5.32}$$

In terms of a station i on route r, $\beta_{\pi_k,i} \cdot d_k$ denotes the expected number of passengers in group k that takes route r and alights at this stop. g^r_i guarantees route r contain this stop and $\hat{y}^r_{k,i}$ ensures passengers choose to alight at at this stop. Therefore, number of passengers $y^r_{k,i}$ chooose to alight at this stop i cannot exceed the expectation value on stop i with maximum benefit for passengers . Together with Constraint (5.32), this set of constraints indicate that passengers in each group taking each route will only choose the best station to alight.

2. **Set upper bound for served passengers according to bus deployment**: Amount of passengers taking a route is further bounded by the deployment issues:

$$\sum_{k \in K, i \in N_r} y^r_{k,i} \leq f_r \cdot Q \;\; \forall r \in R \tag{5.33}$$

In Constraint (5.33), Q denotes the bus capacity and this constraint guarantees that the number of passengers served by bus cannot exceed total bus capacities provided by the operators.

3. **Set upper bound for deployed buses**: In real-world operations, the number of buses deployed on a route must be within a reasonable range, otherwise traffic jams may occur. The following constraint limits the number of buses

on each route r not exceed the maximum number of buses F.

$$f^r \leq F \ \ \forall r \in R \tag{5.34}$$

$$f^r \leq \frac{T}{\delta^r} \cdot h^r \ \ \forall r \in R \tag{5.35}$$

Constraint (5.35) ensures the provided services must be limit by the number of buses offered by operators. T is the total time horizon and the round of services for a particular route is bounded by the travel time on this route. Shorter travel time allows a bus provide multiple round services along a route r. The more buses offered for a route x_r, the more round services can be provided accordingly.

4. **Demand constraint**: The following constraint ensures the served amount of passengers plus the unserved amount of passengers equal to the total number of passengers. d_k is the demand for group k.

$$\sum_{r \in R, i \in N_r} y^r_{k,i} + \bar{y}_k = d_k \ \ \forall k \in K \tag{5.36}$$

5. **Budget constraint**: In this context, operators' cost is mainly from two aspects: fixed cost of offering buses and variable cost of operations. We use the following constraints to represent the cost scheme and hence set constraints on the cost by the total budget.

$$\sum_{r \in R} c_x \cdot h^r + \sum_{r \in R} c_f \cdot f^r \leq B_1 \tag{5.37}$$

Parameter c_h denotes the cost related to buying or renting a new route and c_q denotes the cost of adding extra round of bus service on the route. The summation of the overall cost is bounded by the total budget B_1.

To help readers recap, we summarize the parameters utilized in this section in Table 5.3.

Table 5.3: Notations for Designing Bus Service

Parameters	Explanations
K	Passenger set, contain demand d_k and destination π_k for each group k
N_r	Node set containing nodes in route r
w_k^r	Whether demand k is covered by route r
g_i^r	Whether station i is covered by route r
Q	Bus capacity
F	Maximum round of bus service for a route

We summarize the model in Table 5.4.

Table 5.4: MILP Model for Bus Service

$$\textbf{min} \sum_{r\in R, i\in N_r, k\in K} (\delta_i^r + \bar{\delta}_{i,\pi_k}) \cdot y_{k,i}^r + \sum_{k\in K} \bar{\delta}_{s,\pi_k} \cdot \bar{y}_k$$

$$y_{k,i}^r \leq \beta_{\pi_k,i} \cdot d_k \cdot g_i^r \cdot \hat{y}_{k,i}^r \quad \forall r \in R, i \in N_r, k \in K \tag{5.38}$$

$$\sum_{k\in K, i\in N_r} y_{k,i}^r \leq f_r \cdot Q \quad \forall r \in R \tag{5.39}$$

$$f^r \leq \frac{T}{\delta^r} \cdot h^r \quad \forall r \in R \tag{5.40}$$

$$f^r \leq F_{max} \quad \forall r \in R \tag{5.41}$$

$$\sum_{r\in R, i\in N_r} y_{k,i}^r + \bar{y}_k = d_k \quad \forall k \in K \tag{5.42}$$

$$\sum_{i\in N_r} \hat{y}_{k,i}^r \leq 1 \quad \forall r \in R, k \in K \tag{5.43}$$

$$\sum_{r\in R} c_x \cdot h^r + \sum_{r\in R} c_f \cdot f^r \leq B_1 \tag{5.44}$$

5.3 Designing Bike Sharing Services

Bus is not the only mode that can distributing passengers over the spatial horizon in public transportation domain. Bike sharing services play an important role in recent years. Unlike bus services, infrastructures for bike services inside the park is not established and we need to plan the location of bike stations. We simplify the model from [15] to design the bike service system.

This model combines strategic decisions for both locating bike stations and defining the operation issues of the system. The bike-sharing model determines the location of the stations, the number of bicycles that should be purchased in each station and the fleet size simultaneously.

The criterion to determine whether such bike sharing system works well is the total travel time achieved by this system. A common issue related to bike service design is the unbalanced inventory caused by the passengers' time-varying demand. We therefore, consider time-dependent travel demand. We introduce decision variable $x^{\tau}_{i,j} \in \mathbb{N}^{+}$ to represent number of passengers moving from i to j at time τ. $\bar{x}_k$ represent the number of passengers that could not take bike services in group k. Therefore, the objective function for our model is to minimize the total travel time for all passengers (5.45).

$$\textbf{minimize} \sum_{k\in K,\tau\in T} \delta_{s,\pi_k} \cdot x^{\tau}_{s,\pi_k} + \bar{\delta}_{s,\pi_k} \cdot \bar{x}_k \tag{5.45}$$

δ_{s,π_k} denotes the travel time associated with riding bikes and $\bar{\delta}_{s,\pi_k}$ represents the relative walking time. The objective function provides optimal design of a bike-sharing system to minimize passengers' travel time while taking the service level into account.

We address the concerns of this problem using following constraints:

1. **Determine the location and capacity of bike stations**: With a set of candidate bike station set, the natural next step is to determine the location and capacity of each candidate bike station.

$$q_i \geq \underline{q} \cdot l_i \;\; \forall i \in N \tag{5.46}$$

$$q_i \leq \bar{q} \cdot l_i \;\; \forall i \in N \tag{5.47}$$

Decision variables $l_i \in \{0, 1\}$ represent whether to choose operate the candidate bike station i or not. Relatively, $q_i \in \mathbb{N}$ denotes the capacity of the bike station. Only when a station is utilized, i.e. $l_i = 1$, the capacity for this station makes sense. Constraint (5.46) gives an upper and lower bound for the real capacity designed.

2. **Balance and rebalance issues**: When passengers travel with bike, they simply put bikes at the stations near their destinations. Overtime, the inventory of bikes one station will be extremely larger than the others. Therefore, operators need to use a truck or whatever methods to rebalance the bike inventory.

$$
\begin{aligned}
& v_i^\tau = v_i^{\tau-1} - \sum_{j\in N} x_{i,j}^{\tau-1} + \sum_{j\in N} x_{j,i}^{\tau-1} \\
& - \sum_{j\in N} r_{i,j}^{\tau-1} + \sum_{j\in N} r_{j,i}^{\tau-1} \quad \forall i \in N, \tau \in T
\end{aligned} \tag{5.48}
$$

$$
\sum_{j\in N} r_{i,j}^\tau \leq v_i^\tau \quad \forall i \in N, \tau \in T \tag{5.49}
$$

$$
v_i^\tau \leq q_i \quad \forall i \in N \tag{5.50}
$$

$$
v_i^0 = v_i^T \quad \forall i \in N \tag{5.51}
$$

We denote $r_{i,j}^\tau \in \mathbb{N}$ as the number of bikes relocated from i to j at time τ. $v_i^\tau \in \mathbb{N}$ represents the number of bikes at bike station i at the beginning of time τ. Constraint (5.48) defines the equality relationship on the number of bicycles at station i between time step τ and time step $\tau - 1$. Concretely, the number of bicycles available at station i time τ, equals to: the number of bicycles available at $\tau - 1$ plus the bicycles relocated to station i, minus the bicycles relocated from station i, as well as taking considerations for those bicycles that are utilized for passengers' movement. Constraint (5.49) tells the reality that the number of bicycles to be relocated at any station i must be lower than the real number of bicycles in the station at any time τ. Constraint (5.50) indicates that the storage of bicycles for any station need to be constrained by the real capacity of the stations. Constraint (5.51) assumes that the number of bicycles at the beginning of the day is the same with the end of the day.

3. **Passengers movement constraint**: Number of passengers moving over the bike sharing system is limited by the state of the stations. If one of the stations between passengers orgin and destination is closed, passengers cannot place

their bikes and thus they cannot travel on this link.

$$x^{\tau}_{s,\pi_k} \leq d_k \cdot l_s \ \ \forall k \in K, \tau \in T \tag{5.52}$$

$$x^{\tau}_{s,\pi_k} \leq d_k \cdot l_{\pi_k} \ \ \forall k \in K, \tau \in T \tag{5.53}$$

$$l_i \leq \sum_{j,\tau} x^{\tau}_{i,j} + x^{\tau}_{j,i} \ \ \forall i \in N \tag{5.54}$$

$$v^{\tau}_i \geq \sum_{j} x^{\tau}_{i,j} \ \ \forall i \in N, \tau \in T \tag{5.55}$$

Constraint (5.52), (5.53) together ensure station at origin s and destination π_k must open if passengers in group k travel on this link. Conversely, if there is no passengers travel into or out of station i, station i can be closed according to Constraint (5.54). Constraint (5.55) is the requirement from designer that the stored number of bicycles must fulfill the passengers' travel demand.

4. **Demand constraint**: The following constraint ensures the served amount of passengers plus the unserved amount of passengers equal to the total number of passengers and the total number of passengers taking bikes cannot exceed the expected value.

$$\bar{x}_k + \sum_{\tau \in T} x^{\tau}_{s,\pi_k} = d_k \ \ \forall k \in K \tag{5.56}$$

$$\sum_{\tau \in T} x^{\tau}_{s,\pi_k} \leq \beta_{s,\pi_k} \cdot d_k \ \ \forall k \in K \tag{5.57}$$

5. **Budget constraint**: The cost of bike sharing system consists of 2 major parts: fixed cost and operational cost. The fixed cost consists of the spending on establishing a bike station, capacity of this station and purchasing bikes. Operational cost mainly refers to the rebalanced fees.

$$\sum_{i \in N} c_q \cdot q_i + \sum_{i \in N} c_l \cdot l_i + \sum_{i,j \in N, \tau \in T} c_r \cdot r^{\tau}_{i,j} + V \cdot c_b \leq B_2 \tag{5.58}$$

Constraint (5.58) has 4 relative terms, including the cost of (1) establishing the stations and designing the capacities; (2) purchasing all bicycles; and (3) re-balancing the bikes. The summation of these cost cannot not exceed the total budget.

To help readers to recap, we list some parameters in table 5.5.

Table 5.5: Notations for Designing Bus Service

Parameters	**Explanations**
$\underline{q}, \bar{q}$	min/max bike station capacity
c_q	cost of designing the capacity of a station
c_l	cost of constructing bike station
c_r	cost of relocation
c_b	cost of purchasing a single bike

We summarize the model in Table 5.6.

5.4 Lagrangian Relaxation

We have two phases of our problem, the first phase is to determine the assignment of portfolio over each transportation mode mode. Hence, we have the total cost not exceeding the total budget:

$$\left[\sum_i c_q \cdot q_i + c_l \cdot l_i + \sum_{j,\tau} c_r \cdot r_{i,j}^{\tau} + V \cdot c_b\right] + \left[\sum_r c_h \cdot h^r + c_f \cdot f^r\right] \leq B \quad (5.72)$$

The second phase is to run each sub model and obtain the optimal deployment under budget constraints. In this phase, we must ensure the total number of served passengers plus the total number of unserved passengers equal to the total demand in each group:

$$\sum_{r,i} y_{k,i}^{r} + \sum_{\tau} x_{s,\pi_k}^{\tau} + \bar{s}_k = d_k \;\; \forall k \quad (5.73)$$

where we use variable $\bar{s}_k$ to represent the passengers that can neither be served by bus or bikes.

The optimal solution is to achive minimal travel time for all passengers including those who take buses, bikes and walking. Therefore, we have the integrated

Table 5.6: MILP Model for Bike Service

$$\textbf{min} \sum_{k\in K,\tau\in T} \delta_{s,\pi_k}\cdot x^{\tau}_{s,\pi_k} + \bar{\delta}_{s,\pi_k}\cdot \bar{x}_k$$

$$q_i \geq \underline{q}\cdot l_i \;\; \forall i \in N \tag{5.59}$$

$$q_i \leq \bar{q}\cdot l_i \;\; \forall i \in N \tag{5.60}$$

$$v_i^{\tau} = v_i^{\tau-1} - \sum_{j\in N} x_{i,j}^{\tau-1} + \sum_{j\in N} x_{j,i}^{\tau-1} - \sum_{j\in N} r_{i,j}^{\tau-1} + \sum_{j\in N} r_{j,i}^{\tau-1} \;\; \forall i \in N, \tau\in T \tag{5.61}$$

$$\sum_{j\in N} r_{i,j}^{\tau} \leq v_i^{\tau} \;\; \forall i\in N, \tau \in T \tag{5.62}$$

$$v_i^{\tau} \leq q_i \;\; \forall i \in N \tag{5.63}$$

$$v_i^{0} = v_i^{T} \;\; \forall i \in N \tag{5.64}$$

$$x^{\tau}_{s,\pi_k} \leq d_k \cdot l_s \;\; \forall k \in K, \tau \in T \tag{5.65}$$

$$x^{\tau}_{s,\pi_k} \leq d_k \cdot l_{\pi_k} \;\; \forall k \in K, \tau \in T \tag{5.66}$$

$$l_i \leq \sum_{j\in N,\tau\in T} x_{i,j}^{\tau} + x_{j,i}^{\tau} \;\; \forall i \in N \tag{5.67}$$

$$v_i^{\tau} \geq \sum_{j\in N} x_{i,j}^{\tau} \;\; \forall i \in N, \tau\in T \tag{5.68}$$

$$\bar{x}_k + \sum_{\tau\in T} x^{\tau}_{s,\pi_k} = d_k \;\; \forall k \in K \tag{5.69}$$

$$\sum_{\tau\in T} x^{\tau}_{s,\pi_k} \leq \beta_{s,\pi_k}\cdot d_k \;\; \forall k \in K \tag{5.70}$$

$$\sum_{i\in N} c_q\cdot q_i + \sum_{i\in N} c_l \cdot l_i + \sum_{i,j\in N,\tau} c_r\cdot r^{\tau}_{i,j} + V\cdot c_b \leq B_2 \tag{5.71}$$

objective function:

$$\min \sum_{k,r,i} (\delta_i^r + \bar{\delta}_{i,\pi_k})\cdot y^r_{k,i} + \sum_{k,t} \delta_{s,\pi_k}\cdot x^{\tau}_{s,\pi_k} + \sum_k \bar{\delta}_{s,\pi_k}\cdot \bar{s}_k \tag{5.74}$$

As the two models are independent in nature, we need to construct them using a systematic approach. Therefore, we adopt Lagrangian Relaxation to accomplish the above task. We exploit the dependencies between these two problem as follows:

Observation 5.4.1 *In bus and bike models, we have:*

- *Variables h^r and f^r capture the cost of bus deployment plan;*

- *Variables q_i, l_i, V and $r^{\tau}_{i,j}$ capture the cost for bike deployment plan.*

These two set of variables only interact in Constraint (5.72) *when combining bus and bike models.*

Observation 5.4.2 *In bus and bike models, we have:*

- *Variables* $y^r_{k,i}$ *capture number of passengers served with bus service;*
- *Variables* x^τ_{s,π_k} *capture number of passengers served with bike service.*

These two set of variables only interact in Constraint (5.73) *when combining bus and bike models.*

Given the above observations, we employ Lagrangian Relaxation and Sub-gradient method [14] to dualize constraints (5.72) and (5.73) using price variables $\lambda, \boldsymbol{\alpha}$. We obtain the following equations:

$$
\begin{aligned}
\mathcal{L}(\lambda,\boldsymbol{\alpha}) \;=\; & \min\Bigg[\sum_{k,r,i}(\delta^r_i+\bar{\delta}_{i,\pi_k})\cdot y^r_{k,i}+\sum_{k,t}\delta_{s,\pi_k}\cdot x^\tau_{s,\pi_k}+\sum_k\bar{\delta}_{s,\pi_k}\cdot\bar{s}_k \\
& +\lambda\Bigg[(\sum_r c_h\cdot h^r+c_f\cdot f^r) \qquad\qquad (5.75)\\
& +(\sum_i c_q\cdot q_i+c_l\cdot l_i+\sum_{j,\tau}c_r\cdot r^\tau_{i,j}+V\cdot c_b)-B\Bigg] \\
& +\sum_k\alpha_k\Bigg[d_k-(\sum_{r,i}y^r_{k,i}+\sum_\tau x^\tau_{s,\pi_k}+\bar{s}_k)\Bigg]\Bigg] \\
=\; & \min\Bigg[\sum_{k,t}\delta_{s,\pi_k}\cdot x^\tau_{s,\pi_k}+\sum_k\bar{\delta}_{s,\pi_k}\cdot\bar{s}_k \\
& +\lambda\cdot(\sum_i c_q\cdot q_i+c_l\cdot l_i+\sum_{j,\tau}c_r\cdot r^\tau_{i,j}+V\cdot c_b-B) \\
& +\sum_k\alpha_k(d_k-\sum_\tau x^\tau_{s,\pi_k}-\bar{s}_k)\Bigg] \\
& +\min\Bigg[\sum_{k,r,i}(\delta^r_i+\bar{\delta}_{i,\pi_k})\cdot y^r_{k,i}+\lambda\cdot(\sum_r c_h\cdot h^r+c_f\cdot f^r) \\
& -\sum_{k,r,i}\alpha_k\cdot y^r_{k,i}\Bigg] \qquad\qquad (5.76)
\end{aligned}
$$

In Equation (5.76), the first part corresponds to the bike problem and the second part corresponds to the bus deployment problem. Specifically, we summary the two slave problems in Table 5.7 and Table 5.8.

$$
\begin{aligned}
\min \quad & \sum_{k,\tau} \delta_k \cdot x^{\tau}_{s,\pi_k} + \sum_{k} \bar{\delta}_{s,\pi_k} \cdot \bar{s}_k \\
& + \lambda \cdot (\sum_{i} c_q \cdot q_i + \sum_{i} c_l \cdot l_i + \sum_{i,j,\tau} c_r \cdot r^{\tau}_{i,j} + V \cdot c_b - B) \\
& + \sum_{k} \alpha_k (d_k - \sum_{\tau} x^{\tau}_{s,\pi_k} - \bar{s}_k) \\
s.t. \quad & Constraint\ (5.48) \ -- \ (5.58) \ \ hold
\end{aligned}
$$

Table 5.7: Solve Bike Deployment

$$
\begin{aligned}
\max \quad & \sum_{k,r,i} (\delta^r_i + \bar{\delta}_{i,\pi_k}) \cdot y^r_{k,i} + \lambda \cdot (\sum_{r} c_x \cdot x^r + \sum_{r} c_f \cdot f^r) - \sum_{k,r,i} \alpha_k \cdot y^r_{k,i} \\
s.t. \quad & Constraint\ (5.31) \ -- \ (5.36) \ \ hold
\end{aligned}
$$

Table 5.8: Solve Bus Deployment

In Lagrangian Relaxation procedure, the dual problem $\min_{\lambda,\boldsymbol{\alpha}} L(\lambda, \boldsymbol{\alpha})$ is solved iteratively with sub-gradient approach. In each iteration, we update λ by:

$$
\lambda^{iter+1} = \left[\lambda^{iter} + \gamma^{iter} (\sum_{r} c_x \cdot x^r + c_f \cdot f^r + \sum_{i} c_q \cdot q_i + c_l \cdot l_i + \sum_{j,\tau} c_r \cdot r^{\tau}_{i,j} + V \cdot c_b - B) \right]_+
$$

$[\cdot]_+$ is to ensure λ always obtain positive value, otherwise we take it 0. Similarly, we update $\boldsymbol{\alpha}$ by:

$$
\alpha^{iter+1}_k = \left[\alpha^{iter}_k + \gamma^{iter} (\sum_{r,i} y^r_{k,i} + \sum_{\tau} x^{\tau}_{s,\pi_k} + \bar{s}_k - d_k) \right]_+ \quad \forall k \tag{5.77}
$$

If the difference between dual objective and primal objective is less than a threshold η, the convergence is detected. We extract the Primal problem by deploying bus service in the first place and utilize the remaining budget B' to deploy bike services. Given solutions X^r and F^r provided by bus deployment in Table 5.8, we have remaining budget:

$$
B' = B - (\sum_{r} c_x \cdot X^r + c_f \cdot F^r) \tag{5.78}
$$

Therefore, the detailed Primal problem is described in Table 5.9.

$$
\begin{aligned}
\max \quad & \sum_{k,\tau} \delta_{s,\pi_k} \cdot x^{\tau}_{s,\pi_k} + \sum_{k} \bar{\delta}_{s,\pi_k} \cdot \bar{s}_k \\
s.t. \quad & Constraint\ (5.46) -- (5.58)\ \ hold \\
& \sum_{i} c_q \cdot q_i + \sum_{i} c_l \cdot l_i + \sum_{i,j,\tau} c_r \cdot r^{\tau}_{i,j} + V \cdot c_b \leq B'
\end{aligned}
$$

Table 5.9: Extract Primal

We summarize the Lagrangian Relaxation process in Algorithm 4. For simplicity, we set:

$$
\begin{aligned}
\mathbb{B} &= \sum_{r} c_x \cdot x^r + c_f \cdot f^r + \sum_{i} c_q \cdot q_i + \sum_{i} c_l \cdot l_i + \sum_{j,\tau} c_r \cdot r^{\tau}_{i,j} + V \cdot c_b - B \\
\mathbb{D} &= \sum_{r,i} y^{r}_{k,i} + \sum_{\tau} x^{\tau}_{s,\pi_k} + \bar{s}_k - d_k
\end{aligned} \tag{5.79}
$$

Algorithm 4 Lagrangian Relaxation Procedure

procedure SOLVELR
  $\lambda \leftarrow \lambda^0, \alpha_k \leftarrow \alpha^0_k, iter \leftarrow 0$
  repeat
    $o_1, \boldsymbol{y}, \boldsymbol{f}, \boldsymbol{h} \leftarrow$ SolveBus $(\lambda^{iter}, \boldsymbol{\alpha}^{iter})$
    $o_2, \mathbf{x} \leftarrow$ SolveBikes $(\lambda^{iter}, \boldsymbol{\alpha}^{iter})$
    $\lambda^{iter+1} \leftarrow [\lambda^{iter} + \gamma^{iter} \cdot \mathbb{B}]_+$
    $\alpha^{iter+1}_k \leftarrow [\alpha^{iter}_k + \gamma^{iter} \cdot \mathbb{D}]_+ \quad \forall k$
    $o_p, \mathbf{x} \leftarrow$ExactPrimal$(\boldsymbol{y}, \boldsymbol{f}, \boldsymbol{h})$
    $\gamma^{iter} \leftarrow \frac{1.2 * \left[o_p - (o_0 + o_1)\right]}{||\mathbb{B}||^2 + ||\mathbb{D}||^2}$
    $iter \leftarrow iter + 1$
  until $o_p - (o_1 + o_2) \leq \delta$
return $o_p, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{f}, \boldsymbol{h}$

5.5 Experiment

We conduct experiment with Java, CPLEX 12.51 and executed on a Core(TM) i7-6700 CPU 3.4GHz processor under Windows.

5.5.1 Dataset

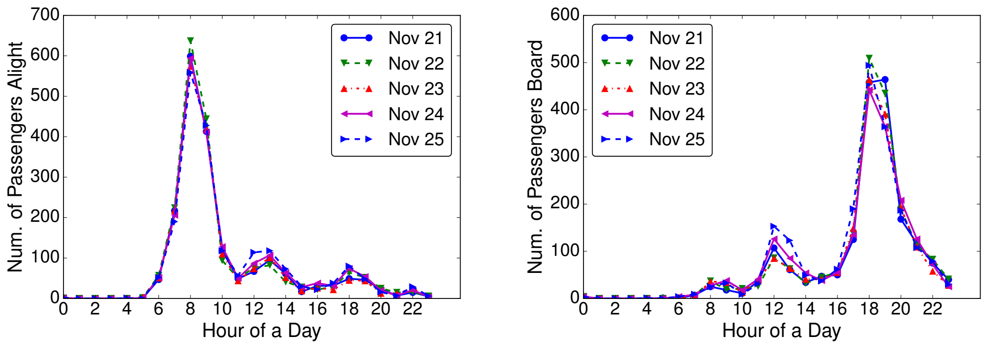
We focus on a business park called One-North in Singapore, which is designed to host a cluster of world-class research facilities and business park space [1]. There is one single MRT station called One-North MRT station serving the hubs of this business district. Most of the passengers alight at One-North MRT station and find their way towards their destinations in the business park. Therefore, we need to effectively distribute the demand particularly during morning and evening peak hours. We concern about providing efficient transportation services that connect One-North MRT station and the office buildings inside the business park. The topology of the transport network is extracted from open street map.

We derive the demand data using the record from transportation fare card, which is called EZLink card in Singapore. EZlink card is mainly used for transportation payment in Singapore. Passengers need to tap their card whenever boarding and alighting. In other words, a full set of information including passengers' boarding time/location, alighting time/location, transportation modes, etc. will be created for each particular card.

The total demand at this business district is extracted from a full week of EZLink data recorded (from 21, November 2011, Monday to 25, November 2011, Friday). From Figure 5.3 we clearly observe two demand peak during morning and evening peak hours.

We adopt the public information provided by JTC webpage [2] to measure the distribution of demand. From which, we obtain the location of office buildings at the business park. One assumption we made here is to assume the number of passengers belongs to each building is proportional to the area occupied by such building. We use Google map to measure the area of each office and use the same ratio to derive the demand for each location.

[1] http://www.jtc.gov.sg/industrial-land-and-space/pages/one-north.aspx
[2] http://landtransportguru.net/web/wp-content/uploads/2016/07/one-north-rider-map-1.jpg



(a) Alighting (b) Boarding

Figure 5.3: Passengers alighting / boarding from One-North MRT station

5.5.2 Penalty for Setting Stations

In real world, most of the time bus runs with a regulate speed yet makes stopping delay at bus stops. Intuitively, bus decelerates when entering a bus station and stop for passengers to board and deboard. After which, it accelerates to normal speed again. Thus stopping delay at bus stop include bus acceleration, deceleration, dwell time [53] and so on. When designing bus lines, we take such delay into considerations by adding up Δ to the travel time for each line.

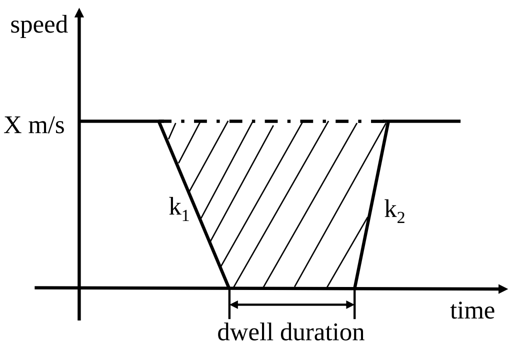


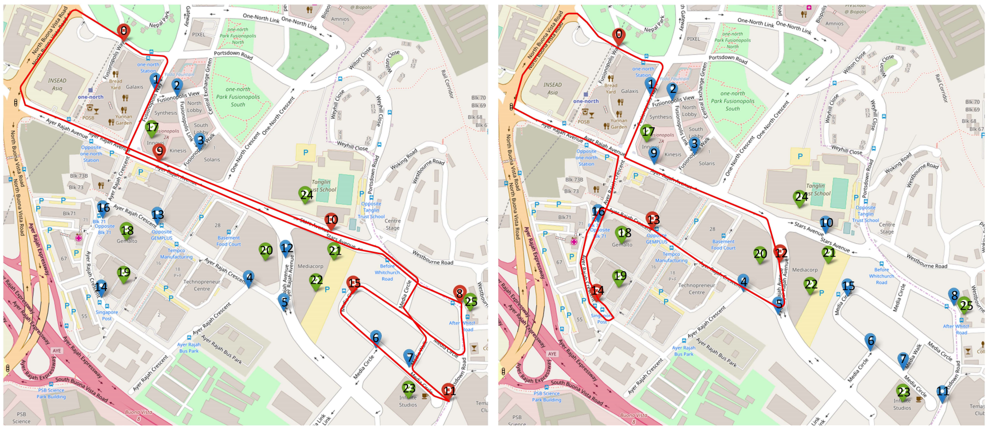
Figure 5.4: Computing the delay for each station

Figure 5.4 is a simple illustration on how to compute the penalty Δ. Initially bus travel with regulate speed which we assume to be $25\ km/h$ according to [11], in other words, $X = 6.94$. When reaching the bus stop, the dwell time is around $30s$ according to [39] and the decelerate slope $k_1 = 1.2$, accelerate $k_2 = 1.2$ based on [53]. The area for the shaded region is the extra distance traveled and $\Delta =$

S_{shaded}/X and thus the delay for each station as $\Delta = 35.78s$.

5.5.3 Results of Column Generation

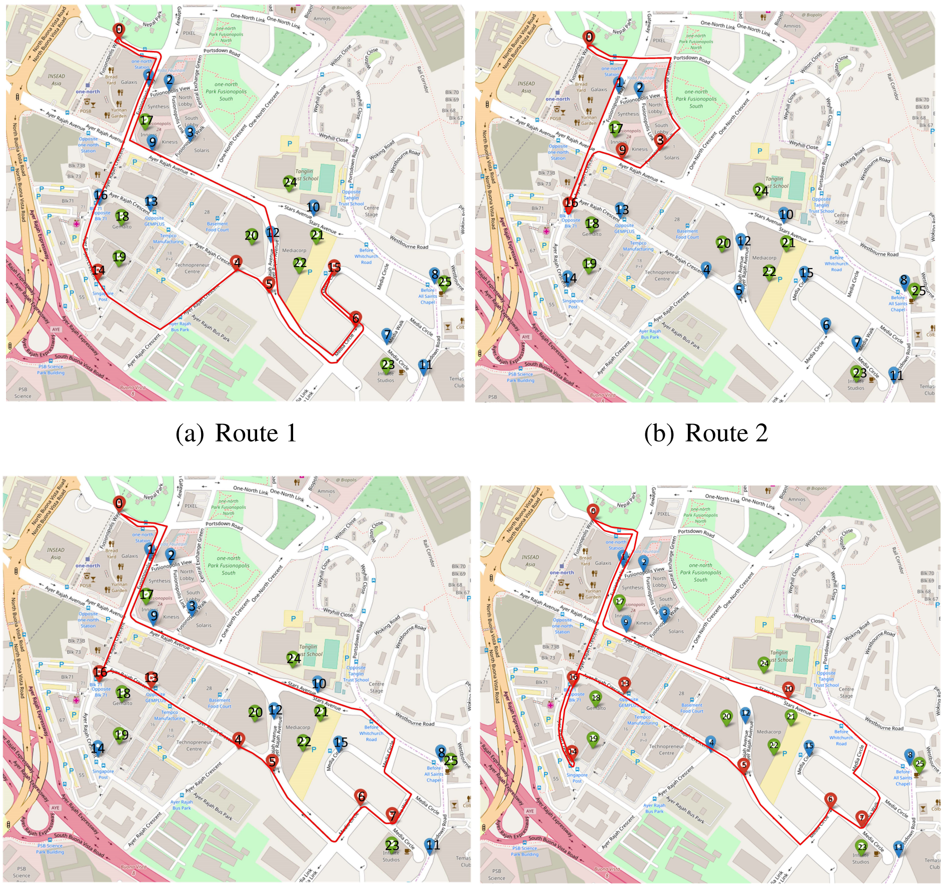
The operators deployed bus service at One-North business park using two routes shown in Figure 5.5 (a) and (b). In this figure, markers with red color are those stations that form a route (red line). Those blue ones denote candidate stations that are not utilized. Passengers' destinations are represented by green markers. Numbers on the markers represent the index of each node.



(a) Route 1 (b) Route 2

Figure 5.5: Existingn routes deployed at One-north business park

We use the two routes that operators provides as the feasible routes and start the column generation procedure to generate new routes. The following routes shown in Figure 5.6 show the routes generated. Routes in Figure 5.6(c) and Figure 5.6(d) look similar at the first glance. However, bus makes stops at different stations in each route. Therefore, they are characterised as different routes. Whether bus makes a stop is due to the penalty of setting stops. Stopping at a station takes extra deceleration and acceleration time, which makes the trip longer. Thus, model chooses those stations that make benefit to the objective function.



(a) Route 1 (b) Route 2

(c) Route 3 (d) Route 4

Figure 5.6: Generated routes

5.5.4 Results of Comparing to Existing Deployment

With the routes generated in the first phase, in this section, we compare the performance of existing settings deployed by operators to our settings. We first discuss the parameter settings in our experiment.

5.5.4.1 Parameter Settings

To set the budget, we amortize the one-year investment and map to the period budget. The annual investment is around 1 million [3]. Our planning horizon is 3 hours as the demand surge occurs during morning peak hours, from $7:00-9:00\ AM$ according to Figure 5.3. The loan interest rate in Singapore is around 1.7%. Therefore, we show the amortization results for annual investment to 3 hour period budget

[3]http://www.straitstimes.com/singapore/transport/big-sponsor-wanted-for-bike-sharing-scheme

in Table 5.10.

Table 5.10: Annual investment and period budget

annual investment	period buget	annual investment	period buget
\$100,000	\$92.10	\$900,000	\$828.93
\$200,000	\$184.21	\$1,000,000	\$921.03
\$300,000	\$276.31	\$1,100,000	\$1,013.14
\$400,000	\$368.41	\$1,200,000	\$1,105.24
\$500,000	\$460.52	\$1,300,000	\$1,197.34
\$600,000	\$552.62	\$1,400,000	\$1,289.45
\$700,000	\$644.72	\$1,500,000	\$1,381.55
\$800,000	\$736.83	\$1,600,000	\$1,473.65

Bus rental fee and bus capacity can be obtained from public information online [4]. In our experiment we take bus with 45 seats and hourly rental fee 85 SGD. In terms of the operational fees, according to a report [5], the daily cost of a bus is around 600 SGD. However, this cost is based on an average trip length 4.3 km over a day [6]. We translate this value to our case where the average trip length is 0.4km and obtain the period operational cost 6.9.

The cost of setting bike services are not published. We thus obtain the information from [15] and translate the Euros to Singapore dollars. In this case, the fixed cost of constructing a station is 4538 SGD. This value is supposed to be a year cost and we divide 1095 to translate to 3 hours time horizon and find out the payment to be 4.15. Following the same way, the cost of per docks in a station is 0.69, per bike 0.41. The relocation fee is 0.02 for each bicycle.

With above settings, the comparison is conducted in two ways: the first comparison is to compare the bus service we provide to the existing service. The second one compares the mixed transportation modes with exiting bus services.

[4]http://www.singaporeluxurylimousine.com/charter-a-bus/
[5]https://www.mot.gov.sg/News-Centre/Highlights/The-$1-1-billion-bus-question/
[6]https://www.lta.gov.sg/content/dam/ltaweb/corp/PublicationsResearch/files /FactsandFigures/Statistics in Brief 2015 FINAL.pdf

5.5.4.2 Comparing Bus Services

We evaluate the performance using passengers average travel time. Intuitively, the lower the travel time, the better the performance. We take the existing routes deployed by operators and the routes generated using column generation to conduct the experiment. As the operators only have 2 routes incorporated in total, we further include an experiment taking as most 2 routes out of all to make a fair comparison.

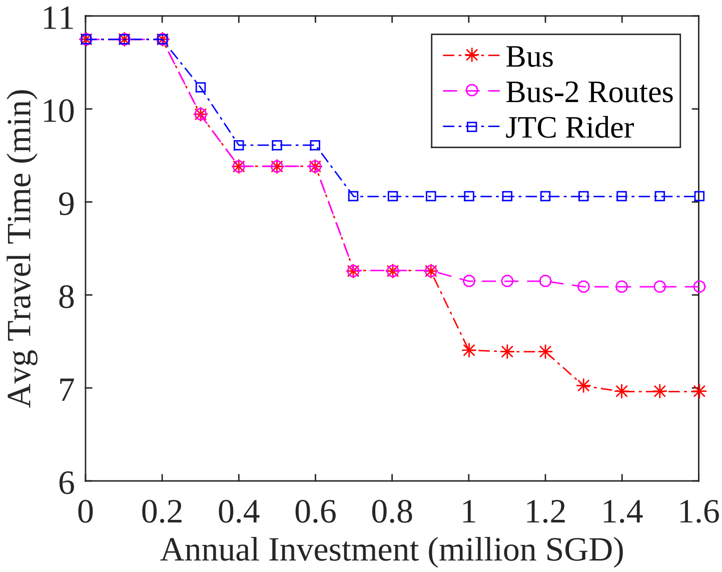
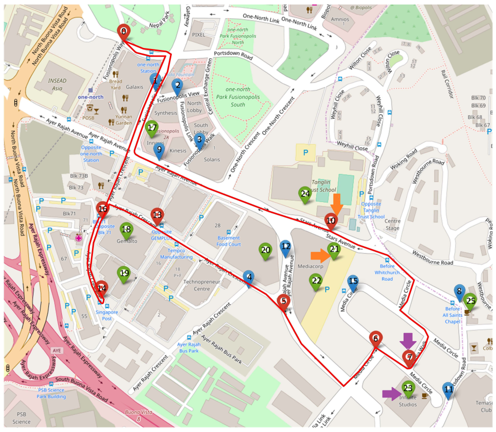


Figure 5.7: Comparison of bus services

Figure 5.7 reports the comparison results with three different services deployed. From this figure, we observe the average travel time does not change when the annual investment is smaller than 0.2 million, since the budget is not enough to implement the bus service at that point. With annual investment increases, travel time of all three modes goes down. This is reasonable as the more budget we invest, the more passengers we can serve with buses. When annual investment is larger than 0.3 million, the average travel time derived by existing JTC Rider is higher than the bus service derived by us. We use the situation of 0.3 million annual investment to explain the result. At this point, the number of passengers served by both modes is the same. Hence, it is the different settings of stops in the route that make a difference.

Figure 5.8 show the bus routes adopted when the annual investment is 0.3 mil-

(a) JTC rider route



(b) Bus service route

Figure 5.8: Bus deployment when annual invest is 0.3

lion for both JTC Rider and our bus service in respective. We mark the stations that passengers alight using arrows. Destinations with the same color represent the group of passengers served by the associated station. In Figure 5.8 (a), 4 groups of passengers with destinations $21, 22, 23$ and 25 are served by station 8. The curiosity occurs when passengers with destination 23 alight from station 8 rather than station 11. This is because if passengers alight at station 11, the total travel time is 11.46 with on board time 8.63 and walk time 2.83. While alighting from station 8 takes total time 6.26 with 4.45 minutes on board and 1.81 minutes walk. Obviously station 8 is a better choice than station 11 for these passengers. In Figure 5.8, 2 groups of passengers with destination 21 and 23 are served by station 10 and 7, receptively.

Figure 5.9 plot the average travel time achieved by different service modes when the annual investment is 0.3 million. X-axis denotes the destination of passengers in various groups, which is consistent with Figure 5.8. This figure explicitly shows how passengers from each destination benefit from the above two services deployed. With no service provided, passengers have to walk towards destinations. Therefore, the travel time associated with no service is an upper bound of travel time for each group of passengers. With JTC Rider and our bus service provided, lower travel time will be achieved. Regarding JTC Rider (represented by the blue bar), the aver-

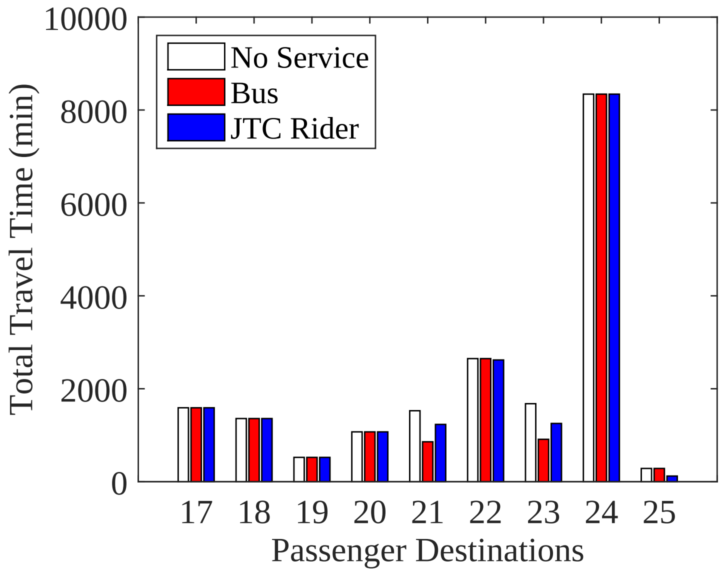


Figure 5.9: Explanation of bus services for annual investment 0.3

age travel time for destination $21, 22, 23$ and 25 is reduced compared to no service mode. However, the reduction is relatively smaller than the reduction achieved by our bus mode. When we take a look at our bus service, the travel time reduction for passengers with destination 21 and 23 is large enough to beat the reduction of JTC Rider. Therefore, at this point when annual investment is 0.3 million, our bus mode achieve lower travel time and from the abovementioned analysis which indicates that our bus route generated by column generation approach is better than the existing one.

Back to Figure 5.7, when the annual investment is 1 million Singapore Dollars, the average travel time for both JTC Rider and bus with 2 routes go flat. This is due to the setting of an upper bound for the resources to assign to each single line. Otherwise too frequent bus services or too many buses running on this line may cause congestions. Therefore, at this point, adding more budget does not help to incorporate the services and thus average travel time remains the same. This is further demonstrated by the Bus service with more than 2 bus routes incorporated. Investing more budget helps devoting resources to a new bus line and thus the average travel time keep going down.

We summarize deployment and bus route when annual investment is 1 million

SGD in Table 5.11. The numbers with bold font represent the stations that passengers choose to alight. Our bus services provide passengers more options to alight. Even with 2 bus, our service can provide 5 good stations rather than 2 compared with existing JTC Rider.

Table 5.11: Explanation of bus services for investment 1.0

Annual Investment	Services	Deployment & Route
1 million	Bus	20 mins: 0-**12**-13-14-0
		25 mins: 0-**10**-**7**-6-**5**-13-16-14-0
		22 mins: 0-**9**-**10**-**8**-**7**-5-4-13-16-14-0
	Bus-2R	20 mins: 0-**12**-13-14-0
		20 mins: 0-**9**-**10**-**8**-**7**-5-4-13-16-14-0
	JTC Rider	20 mins: 0-**12**-13-14-0
		36 mins: 0-**8**-15-11-10-9-0

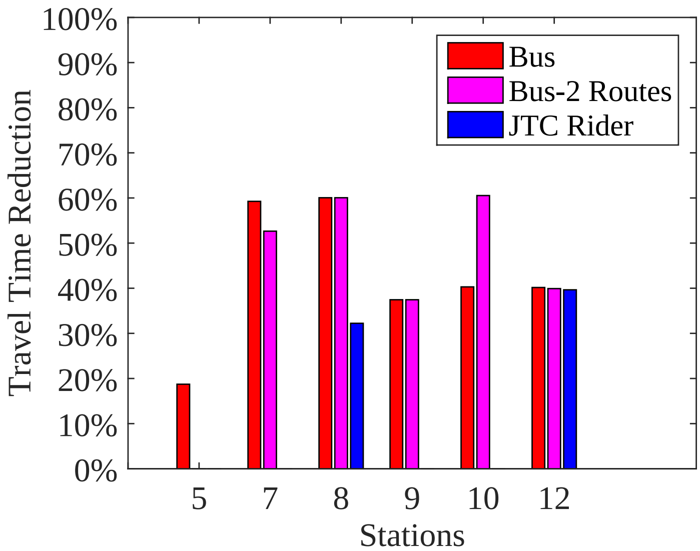


Figure 5.10: Explanation of bus services for annual investment 1.0

Setting different stations offer passengers different travel time reduction. We summarize all possible stations and their relative travel time reduction for 3 different transportation services in Figure 5.10. From this figure, we observe for station 8 and 12 utilized by JTC Rider; the travel time reduction is smaller compared to the bus service provided by our model. Moreover, those bus services can take other stations to make travel time reduction. Therefore, our model achieves smaller travel time than existing deployment at this point.

5.5.4.3 Comparing Mixed Transportation Options

The second comparison is conducted between the bus service and a mixed transportation options of both bus and bike services. Figure 5.11 reports the comparison

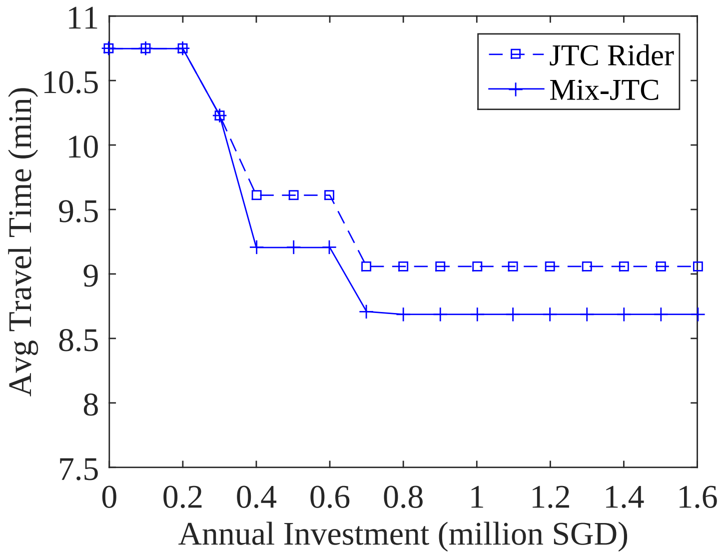


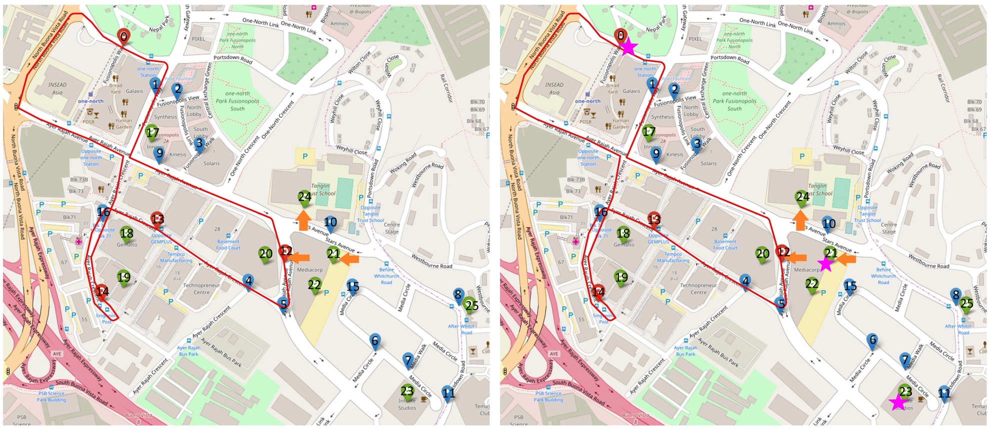
Figure 5.11: Comparison of bimodal options

results between existing bus service and mixed transportation options for JTC Rider. First of all, we take a look at bus model and the mixed service. Before the annual investment reaches 0.4 million, model chooses bus service to reduce the travel time. In this case, the two lines are identical. When the annual investment is 0.4 million, the budget starts to invest to bike service to serve passengers. Details are presented in Table 5.12 where we notice that 57.16 budget is assigned to bike service to serve 55 passengers.

Table 5.12: Budget and passengers distribution with annual investment 0.4

	Budget assignment	Num. of served passengers
JTC Rider	bus: 368.41 bike: 0	bus: 405 bike: 0
JTC+bike	bus: 311.25 bike: 57.16	bus: 405 bike: 55

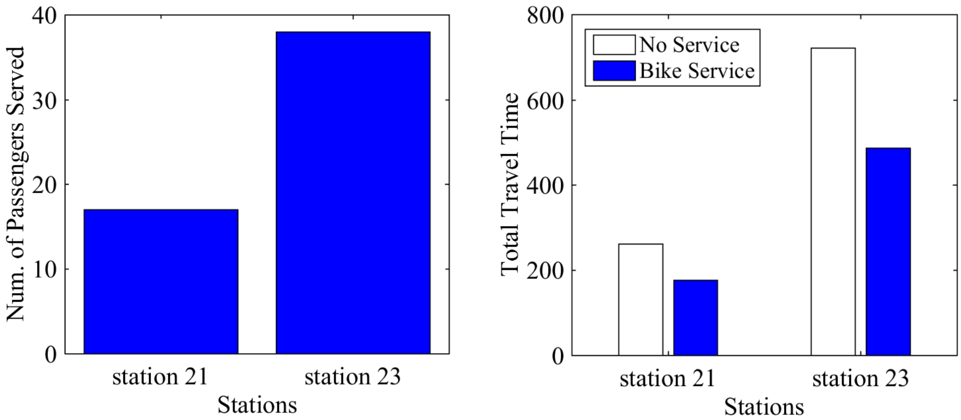
We plot the detailed deployment when annual investment is 0.4 million in Figure 5.12. Figure 5.12 (a) plots the utilized bus route with passengers' choice over alighting stations and Figure 5.12 (b) added the bike station location deployment represented with pink stars.



(a) Route (b) Route and bike station

Figure 5.12: JTC rider deployment

Intuitively, without implementing bike services, passengers with destinations 21 and 24 choose to alight at their nearest station 12. In Figure 5.12 (b), there are 3 bike stations constructed, with capacity $20, 5$ and 15 for stations $0, 21$ and 23 in respective. Constructing a bike station help serve passengers and reduce passengers' travel time. We further plot the travel time achieved by each bike station in Figure 5.13. In Figure 5.13 (a), we observe station 21 and 23 serve 17 and 38 passengers,



(a) Served passengers (b) Travel time

Figure 5.13: Performance of bike service when annual investment is 0.4

in respective. In Figure 5.13 (b), the white bar with no service is the walking time for those un-served passengers. The blue bar is the travel time when riding bike for those served passengers. Providing bike service helps reduce 32.5% and 32.5% travel time reduction for passengers with destination 21 and 23.

When annual investment reaches 1.6 million, JTC Rider maintains 2 of its bus routes and operate the way shown in Table 5.13. Connecting Table 5.13 and 5.14, with bike services, 4 bike stations can be constructed with around 5% of the total budget. 60 passengers are served by bike with the distribution and relative travel time reduction shown in Figure 5.14.

Table 5.13: Explanation of jtc bimodal services for investment 1.6

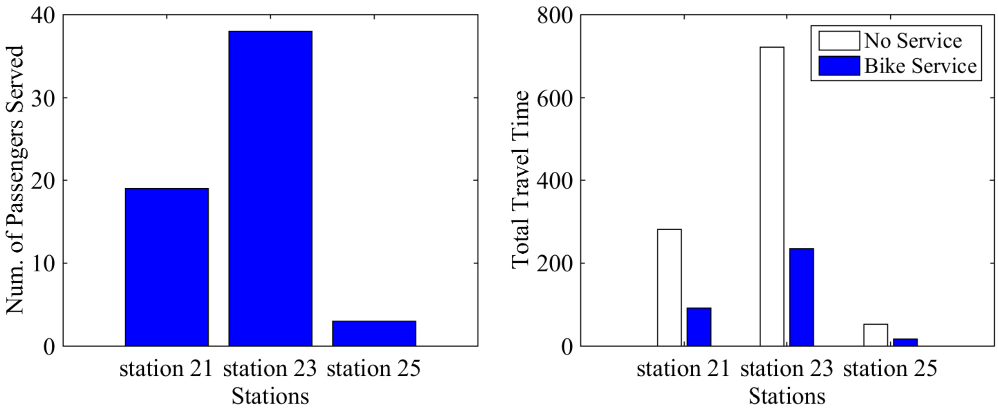
Annual Investment	Modes	Deployment & Route
1.6 million	JTC Rider	20 mins: 0-**12**-13-14-0 36 mins: 0-**8**-15-11-10-9-0
	JTC Rider + Bike	20 mins: 0-**12**-13-14-0 45 mins: 0-**8**-15-11-10-9-0 stations: | 0 | 21 | 23 | 25 | capacity:| 20 | 19 | 20 | 20 |

Table 5.14: Budget and passengers distribution with annual invest 1.6

	Budget assignment	Num. of served passengers
JTC Rider	bus: 1842.07 bike: 0	bus: 604 bike:0
JTC+bike	bus: 1753.76 bike: 88.31	bus: 585 bike: 60

5.6 Conclusion

Managing the demand surges during morning and evening peak hours at business park is a challenging task. Existing strategies utilize a single transportation mode (bus) to serve passengers. We improved this approach by (1). Applying column generation approach to derive better bus operation routes; (2). Adding bike sharing option to help expand the transportation capacity, which can further reduce passengers' travel time. Regarding handle the demand surges, bike sharing service is not



(a) Served passengers

(b) Travel time

Figure 5.14: Performance of bike service when annual investment is 1.6

flexible as altering the existing plan is hard once the infrastructures are set. Bus service gives more flexibility for operators to handle such sudden high demand by renting and deploying more buses. Our work offers a handbook for operators to determine the deployment of bus and bike services in the business park. We validate the plan in the context of a Singapore business park and results indicate our method helps reduce passengers' travel time.

5.7 Discussion

The bike sharing model described in this thesis uses fixed bike stations. As passengers have to take and return bikes at the bike stations, the bike journeys are limited by the number of bikes/docks available at the bike station and the location of the bike stations.

Unlike the traditional bike sharing, the station-less bike sharing scheme works in a different way, which does not require the constructions of bike stations. Each bike is installed with GPS system that allows users to locate them using a map from their smartphone app. Once users reach the destinations, they just park the bicycles at any public bicycle parking area and lock the bike back up using the clamp. Their is no more need for them to park the bike to bike stations.

It is possible for the proposed bike sharing model to be extended to stationless bike sharing scheme with some following modifications. The cost for the station should set to be 0 and the trips are not limited by the availabilities of bikes and docks in stations. The unit price associated with each bike needs to be increased as additional each bike is required to install some additional accessories, such as the GPS sensor. By the way, the cost for developing smartphone app needs to be addressed.

Chapter 6

Conclusions and Future Work

In urban city, most infrastructures are designed to cater a planned capacity, yet surges do happen in times and this has long been a major challenge for urban planners.

In industry, the surging of demand is resolved from the economic perspective where operators use price as a tool to alter the supply and demand. One of the examples is Uber's surging price mechanism. (Uber is a transportation network company who develops, markets and operates the Uber car transportation via mobile apps.) This system justifies the price of ridership based on the demand-supply curve. When detecting the passengers' demand is over-full, Uber starts to raise a multiplier (2x, 3x and so on) on the original price during regular hours to balance the supply-demand relationship.

In this thesis, I presented research on both proactive and reactive strategies to handle the demand surges in urban crowds. Proactive strategy refers to distributing the excessive demand in the temporal dimension. To prevent the demand peaks, this strategy is proposed to provide incentives or distractions such that can slow down the flow rate towards the congested region. Reactive strategy deals with the demand in spatial dimension regarding the situation where it is impossible to delay the crowd movement. For example, after major events, most of the people prefer to go home as soon as possible regardless what incentives are provided. The focus of reactive strategy is to redirect the overflow demand to other parts of the network spatially.

In section 6.1 and 6.2, I review the two strategies and discuss their limitations from both methodological and practical perspectives. Following each limitation, I discuss several lines of future works arising from this dissertation.

6.1 Proactive Strategy

In Chapter 3, we developed model to learn diffusion dynamics parameters and provide decision support to alter the diffusion dynamics to achieve a desired outcome in dependent cascade models.

On the **methodology** development front, existing works at *learning stage* focus on independent diffusions on all outgoing edges of a node. We contribute to incorporate the real-world features (such as learning from aggregate data, modeling queues at network nodes, etc.) into the model, and address flow conservation at network nodes. To account for the lack of data on visitors entering and exiting the theme park as well as taking breaks, we also introduce the "leisure node" with unlimited capacity. Experiment results concretely show how visitors' behaviors can be captured with this node. At the *controlling stage*, we developed an optimization-based approach to compute an optimal plan of management actions to control the underlying diffusion process.

A limitation of this work is to model the transitions as a Hidden Markov Model at the *learning stage*, where the future transitions only depend on the current state and are independent of the history. Though this model is experimentally validated to be valid in our paper, incorporating information from the past may provide more accuracy. To account for this limitation, a possible way is to apply a cutting-edge technique in deep learning – Long Short-Term Memory (LSTM) [20] to learn the model. Typically, LSTM is a Recurrent Neural Network (RNN) architecture which is capable of doing sequential, or time-series learning and predicting. It gets long-term dependencies by connecting previous information to the present task. By appropriately setting the parameters, we can obtain the parameters learned from dependent cascade model using LSTM and compare it to our proposed model.

On the **application-related** research front, we have tested the above-mentioned approach using real-world data from a theme park in Singapore. We have experimentally shown that our learning approach achieves an accuracy close to 80% for popular attractions and the decision support algorithm can provide about $10 - 20\%$ reduction in wait time. Though this model is only demonstrated to be effective in managing the congestion in theme park context, it's capability is not limited in theme park domain.

Another application of this model is to apply the dependent cascade model to learn and control diffusions over transportation network. In Singapore, road cameras record the aggregate number of vehicles appear at a particular road crossing at the different time of a day. From which we can learn the diffusion dynamics over the traffic network. This data set also provides a convenient way to validate transition probabilities in the learned model given the moving behavior of each vehicle recorded. With the learned diffusion models, we can observe the congestions caused by vehicles' movement. One of the strategies to control the congestion is to use Electronic Road Pricing (ERP). ERP is a mechanism deployed by Singapore government to monitor the congestion. Motorists are charged when they use priced roads during peak hours. We can further investigate the best roads to apply ERP tolls to ease the road congestions over the networks.

6.2 Reactive Strategy

In Chapter 4 and Chapter 5, we introduce the usage of public transportation options to disperse the ultra-high demand and in the meantime, minimize passengers' overall travel time.

On the **methodology** development front, we introduce a two-phase optimization based approach for designing bus service, which generates efficient candidate bus routes based on commuters demand and plans the bus deployment such that minimize passengers total travel time. We further extend this service by incorporating bike sharing system to expand the transportation capacity from a central station to passengers' destinations in the context of the business district. In current work, our focus is on capacity planning, where the input demand is static, and the output deployment is an expected outcome over the whole time horizon. Hence, the proposed decision is made to determine the capacity needed by operators to serve the appeared demand. However, for a real-world problem, passengers arrive at the source node in batches.

An extension of the model is to incorporate time-dependent features for the passengers' arrivals. One interesting problem is to capture the time-dependent demand by putting a time index t on demand variable $\boldsymbol{d}$. Based on such time-dependent, origin-to-destination demand records from EZlink data set, an optimization-based model computing and adjusting bus timetables need to be developed. In particular, this model should determine the departure timing for each bus at the source station. The even time space for the departure of each bus may lead to long waiting time for passengers during peak hours. Therefore, an uneven time space schedule for bus service may help accommodate peak-hour demand and in the meantime, maintain some degree of service for passengers during non-peak hours.

On the **application-related** research front, the proposed approach is validated in the context of Singapore community. We extracted passengers' demand information from EZlink data set and conduct the experiment with different settings in Chapter 4: (1). in a normal situation where existing train network is in good operation; (2). in the case of train disruption case where current train system breaks down somewhere. Results show that the optimization approach achieves reduction in the travel time. In Chapter 5, we measured the performance of our approach in business district during morning and evening peak hours. Existing strategies utilize a single transportation mode (bus) to serve passengers. We improved this approach by (1). Applying column generation approach to derive better bus operation routes; (2). Adding bike sharing option to help expand the transportation capacity, which can further reduce passengers' travel time. However, some of the key parameters cannot be obtained, such as the cost of setting bike services, the distribution of passengers' destinations for those attend the concert. In fact, the best way to validate the performance of our model is to use real-world data. To achieve this, we must cooperate with operators and collect empirical data from them. Utilizing such data set, we can estimate the accuracy our model by deploying resources in the real-world and measure the performance of our approaches in reducing the congestion. In this way, the effectiveness can be further validated, and we can adjust our model according to the real-world requirement as well.

Appendix

Non-sub-modularity proof

A function f is sub-modular if: for all $X \subseteq Y$ and $x \in \Omega \backslash Y$, the marginal gain of adding x to the set Y is no more than the gain of adding x to X.

$$f(X \cup x) - f(X) \geq f(Y \cup x) - f(Y)$$

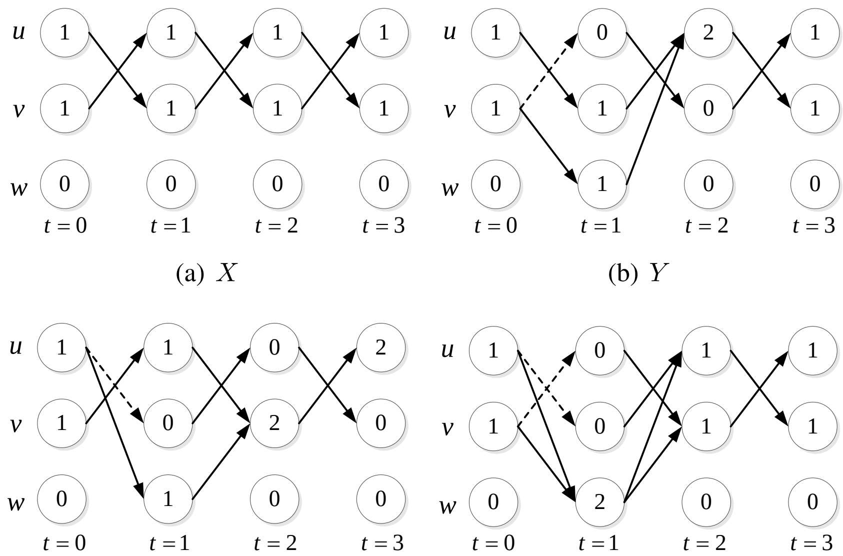
In our problem, Ω is the set of all possible side shows that can be added into the system. Function f corresponds to:

$$-\sum_{u \in \mathcal{A}} \sum_{t \in T} \frac{n_u^t}{s_u} \tag{6.1}$$

We use a counter-example to show that our problem is not sub-modular. Consider there are 3 nodes u, v and w in the system shown in figure (6.1). Node u and v represent the ordinary nodes in the network and w is the buffer node. We set the service rate $s_u = 1$ and $s_v = 2$. Constant transition probabilities $p_{u,v}^t = 1, p_{v,u}^t = 1$ holds for all time steps t. Initially, both $n_u^{t=0}$ and $n_u^{t=0}$ equal to 1.

Empty set X describes the case where there is no side shows taken into the system. Set $x = \{a_{u,v}^0\}$ denotes that at time $t = 0$, a show is taken between node u and v. Similarly, set $Y = \{a_{v,u}^0\}$ shows that a show is added between node v and node u at time 0. In figure (6.1), we explicitly explain the calculation of function f under above scenarios. In this figure, numbers in the circle is n_u^t for each node u.

Solid lines represents transition probabilities $\mathbf{p}$ without any interventions, while the dotted line represents the locations where side shows are taken.



(a) X (b) Y

(c) $X \cup x$ (d) $Y \cup x$

Figure 6.1: Non-Submodularity

Figure (6.1(a)) shows the case where there is no intervention in the system. According to this figure, we calculate $f(X)$ as follows:

$$f(X) = f(\emptyset) = -\sum_{u \in \mathcal{A}} \sum_{t \in T} \frac{n_u^t}{s_u} = -(\frac{4}{1} + \frac{4}{2}) = -6 \tag{6.2}$$

In figure (6.1(b)), transitions with action set Y is described. At $t = 0$, transitions from node v to u is blocked and this flow goes into buffer node w at $t = 0$. After 1 time step delay, the flow starts from buffer node w and transits to its original target node u at $t = 1$. In this case, $f(Y)$ is calculated as:

$$f(Y) = -(\frac{4}{1} + \frac{3}{2}) = -5.5 \tag{6.3}$$

Similarly, calculation of $f(X \cup x)$ and $f(Y \cup x)$ is shown in figure (6.1(c)) and (6.1(d)), respectively. According to the figures, $f(X \cup x) = -5.5$ and $f(Y \cup x) = -4.5$.

From the above examples, $f(X \cup x) - f(X) = 0.5$ and $f(Y \cup x) - f(Y) = 1$.

As $f(X \cup x) - f(X) < f(Y \cup x) - f(Y) = 1$, we conclude that our problem is not sub-modular.

Bibliography

[1] Reza H Ahmadi. Managing capacity and flow at theme parks. *Operations research*, 45(1):1–13, 1997.

[2] Sasan Amini, Eftychios Papapanagiotou, and Fritz Busch. Traffic management for major events. *Digital Mobility Platforms and Ecosystems*, page 187, 2016.

[3] Janett Bttner and Tom Petersen. Optimising bike sharing in european cities - a handbook. *Transport Systems and Logistics*, 2011.

[4] Jinbo Bi Bij, R Edu, and Kristin P Bennett Bennek. Regression error characteristic curves. In *Proceedings of The 20th International Conference on Machine Learning*, pages 43–50, 2003.

[5] Ralf Borndörfer, Martin Grötschel, and Marc E. Pfetsch. A column-generation approach to line planning in public transport. *Transportation Science*, 41(1):123–132, 2007.

[6] Amber Brown, Jacqueline Kappes, and Joe Marks. Mitigating theme park crowding with incentives and information on mobile devices. *Journal of Travel Research*, 52(4):426–436, 2013.

[7] Shih-Fen Cheng, Larry Lin, Jiali Du, Hoong Chuin Lau, and Pradeep Varakantham. An agent-based simulation approach to experience management in theme parks. In *Proceedings of the 2013 Winter Simulation Conference: Simulation: Making Decisions in a Complex World*, pages 1527–1538, 2013.

[8] Carlos F Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, 28(4):269–287, 1994.

[9] Hadi Daneshmand, Manuel Gomez-Rodriguez, Le Song, and Bernhard Schoelkopf. Estimating diffusion network structures: Recovery conditions, sample complexity & soft-thresholding algorithm. In *Proceedings of the 31th International Conference on Machine Learning*, pages 793–801, 2014.

[10] Mauro Dell'Amico, Eleni Hadjicostantinou, Manuel Iori, and Stefano Novellani. The bike sharing rebalancing problem: Mathematical formulations and benchmark instances. *Omega*, 45(2):719, 2013.

[11] Shobhakar Dhakal. De-coupling of urban mobility need from environmental degradation in singapore. Technical report, Working Paper of Urban Environmental Management Project, Institute for Global Environmental Strategies (IGES). Kitakyushu: IGES, 2002.

[12] Anthony Downs. The law of peak-hour expressway congestion. *Traffic Quarterly*, 16(3), 1962.

[13] MaS Elmohamed, Dexter Kozen, and Daniel Sheldon. Collective inference on markov models for modeling bird migration. In *Advances in Neural Information Processing Systems*, pages 1321–1328, 2007.

[14] Marshall L Fisher. The lagrangian relaxation method for solving integer programming problems. *Management science*, 27(1):1–18, 1981.

[15] Ines Frade and Anabela Ribeiro. Bike-sharing stations: A maximal covering location approach. *Transportation Research Part A General*, 82:216–227, 2015.

[16] David K. George and Cathy H. Xia. Fleet-sizing and service availability for a vehicle rental system via closed queueing networks. *European Journal of Operational Research*, 211(1):198–207, 2011.

[17] Amihai Glazer and Esko Niskanen. Parking fees and congestion. *Regional Science and Urban Economics*, 22(1):123–132, 1992.

[18] Manuel Gomez Rodriguez, Jure Leskovec, and Andreas Krause. Inferring networks of diffusion and influence. *ACM Transactions on Knowledge Discovery From Data*, 5(4):21:1–21:37, 2012.

[19] Gabriel Gutiérrez-Jarpa, Carlos Obreque, Gilbert Laporte, and Vladimir Marianov. Rapid transit network design for optimal cost and origin–destination demand capture. *Computers & Operations Research*, 40(12):3000–3009, 2013.

[20] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[21] Kaj Holmberg and Di Yuan. A multicommodity network-flow problem with side constraints on paths solved by column generation. *INFORMS Journal on Computing*, 15(1):42–57, 2003.

[22] Natarajan Janarthanan and Jerry B. Schneider. Computer-aided design as applied to transit system emergency contingency planning. *Computers, Environment and Urban Systems*, 9(1):33–52, 1984.

[23] Jian Gang Jin, Kwong Meng Teo, and Amedeo R Odoni. Optimizing bus bridging services in response to disruptions of urban transit rail networks. *Transportation Science*, 2015.

[24] Jian Gang Jin, Kwong Meng Teo, and Amedeo R Odoni. Optimizing bus bridging services in response to disruptions of urban transit rail networks. *Transportation Science*, 50(3):790–804, 2016.

[25] Mor Kaspi and Tal Raviv. Service-oriented line planning and timetabling for passenger trains. *Transportation Science*, 47(3):295–311, 2013.

[26] Takashi Kataoka, Hidenori Kawamura, Koichi Kurumatani, and Azuma Ohuchi. Distributed visitors coordination system in theme park problem. In *International Workshop on Massively Multiagent Systems*, pages 335–348, 2004.

[27] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 137–146, 2003.

[28] Andreas Krause, Jure Leskovec, Carlos Guestrin, Jeanne VanBriesen, and Christos Faloutsos. Efficient sensor placement optimization for securing large water distribution networks. *Journal of Water Resources Planning and Management*, 134(6):516–526, 2008.

[29] Akshat Kumar, Daniel Sheldon, and Biplav Srivastava. Collective diffusion over networks: Models and inference. In *Proceedings of the 29th Conference on Uncertainty in Artificial Intelligence*, pages 351–360, 2013.

[30] Simon Kwoczek, Sergio Di Martino, and Wolfgang Nejdl. Predicting traffic congestion in presence of planned special events. In *DMS*, pages 357–364, 2014.

[31] Simon Kwoczek, Sergio Di Martino, and Wolfgang Nejdl. Stuck around the stadium? an approach to identify road segments affected by planned special events. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pages 1255–1260. IEEE, 2015.

[32] Suzanne Lassacher, David Veneziano, Steve Albert, and Zhirui Ye. Traffic management of special events in small communities. *Transportation Research Record: Journal of the Transportation Research Board*, (2099):85–93, 2009.

[33] Young-Jae Lee and Vukan R Vuchic. Transit network design with variable demand. *Journal of Transportation Engineering*, 131(1):1–10, 2005.

[34] Jure Leskovec, Lada A Adamic, and Bernardo A Huberman. The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)*, 1(1):5, 2007.

[35] Henry X. Liu, Jeff X Ban, Wenteng Ma, and Pitu B. Mirchandani. Model reference adaptive control framework for real-time traffic management under emergency evacuation. *Journal of urban planning and development*, 133(1):43–50, 2007.

[36] Siyuan Liu, Yisong Yue, and Ramayya Krishnan. Adaptive collective routing using gaussian process dynamic congestion models. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 704–712, 2013.

[37] Ying Liu, Xiaorong Lai, and Gang-Len Chang. Two-level integrated optimization system for planning of emergency evacuation. *Journal of transportation Engineering*, 132(10):800–807, 2006.

[38] Luis M. Martinez, Lus Caetano, Toms Eir, and Francisco Cruz. An optimisation algorithm to establish the location of stations of a mixed fleet biking system: An application to the city of lisbon. *Procedia - Social and Behavioral Sciences*, 54(1):513–524, 2012.

[39] Qiang Meng and Xiaobo Qu. Bus dwell time estimation at bus bays: A probabilistic approach. *Transportation Research Part C: Emerging Technologies*, 36:61–71, 2013.

[40] Michael D Meyer and Peter Belobaba. Contingency planning for response to urban transportation system disruptions. *Journal of the American Planning Association*, 48(4):454–465, 1982.

[41] Seth Myers and Jure Leskovec. On the convexity of latent social network inference. In *Advances in Neural Information Processing Systems*, pages 1741–1749, 2010.

[42] Praneeth Netrapalli and Sujay Sanghavi. Learning the graph of epidemic cascades. In *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems*, pages 211–222, 2012.

[43] Xiaoshan Pan, Charles S Han, Ken Dauber, and Kincho H Law. A multi-agent based framework for the simulation of human and social behaviors during emergency evacuations. *Ai & Society*, 22(2):113–132, 2007.

[44] Brendan Pender, Graham Currie, Alexa Delbosc, and Nirajan Shiwakoti. Disruption recovery in passenger railways. *Transportation Research Record: Journal of the Transportation Research Board*, 2353(1):22–32, 2013.

[45] Brendan Pender, Graham Currie, Alexa Delbosc, and Nirajan Shiwakoti. Disruption recovery in passenger railways: International survey. *Transportation Research Record: Journal of the Transportation Research Board*, (2353):22–32, 2013.

[46] Qiang Qu, Siyuan Liu, Christian S. Jensen, Feida Zhu, and Christos Faloutsos. Interestingness-driven diffusion process summarization in dynamic networks. In *Proceedings of 2014 Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 597–613. Springer, 2014.

[47] Anita Schöbel. Line planning in public transportation: models and methods. *OR Spectrum*, 34(3):491–510, 2012.

[48] Anita Schöbel and Susanne Scholl. Line planning with minimal traveling time. In *2005-5th Workshop on Algorithmic Methods and Models for Optimization of Railways*, 2006.

[49] Susan Shaheen, Stacey Guzman, and Hua Zhang. Bikesharing in europe, the americas, and asia: past, present, and future. *Transportation Research Record: Journal of the Transportation Research Board*, (2143):159–167, 2010.

[50] Daniel Sheldon, Bistra Dilkina, Adam N. Elmachtoub, Ryan Finseth, Ashish Sabharwal, Jon Conrad, Carla P Gomes, David Shmoys, William Allen, Ole Amundsen, et al. Maximizing the spread of cascades using network design. In *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence*, pages 517–526, 2010.

[51] Daniel Sheldon, Tao Sun, Akshat Kumar, and Thomas G Dietterich. Proceedings of the 30th international conference on machine learning. In *International Conference on Machine Learning*, pages 1004–1012, May 2013.

[52] Kut C. So and Christopher S. Tang. On managing operating capacity to reduce congestion in service systems. *European Journal of Operational Research*, 92(1):83–98, 1996.

[53] Alejandro Tirachini. The economics and engineering of bus stops: Spacing, design and congestion. *Transportation research part A: policy and practice*, 59:37–57, 2014.

[54] Jason Tsai, Emma Bowring, Stacy Marsella, and Milind Tambe. Empirical evaluation of computational fear contagion models in crowd dispersions. *Autonomous Agents and Multi-agent Systems*, 27(2):200–217, 2013.

[55] Jeffrey Tubbs and Brian Meacham. *Egress design solutions: A guide to evacuation and crowd management planning*. John Wiley & Sons, 2007.

[56] Evelien Van der Hurk, Haris N Koutsopoulos, Nigel Wilson, Leo Kroon, and Gábor Maróti. Shuttle planning for link closures in urban public transport networks. *Transportation Science*, 2016.

[57] François Vanderbeck. *Decomposition and column generation for integer programs*. 1994.

[58] Vogel and Patrick. *Service Network Design of Bike Sharing Systems*. Springer International Publishing, 2016.

[59] Liaoruo Wang, Stefano Ermon, and John E Hopcroft. Feature-enhanced probabilistic models for diffusion network inference. In *Proceedings of the 2012 European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 499–514, 2012.

[60] Xiaojian Wu, Akshat Kumar, Daniel Sheldon, and Shlomo Zilberstein. Parameter learning for latent network diffusion. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence*, pages 2923–2930, 2013.

[61] Hai Yan and William HK Lam. Optimal road tolls under conditions of queueing and congestion. *Transportation Research Part A: Policy and Practice*, 30(5):319–332, 1996.